

# THIS WEEK

## EDITORIALS

**CLIMATE** Berkeley study generates more heat than light **p.428**

**WORLD VIEW** The true challenge of California's clean-energy pledge **p.429**

**BIRD BRAINS** Tool use not just for the best and brightest **p.431**



## A priceless resource

*The key to treatments for autism and schizophrenia could lie in the brains of recently deceased children. To make advances, researchers need access to an international bank of donated material.*

Few subjects in modern science are as emotive as research on the brains of recently deceased children. The pay-off may seem vague — progress towards understanding, and perhaps treating, neurological conditions including autism and schizophrenia. The difficulties, however, are clearer — the brains must be donated by grieving parents who have just lost a child in sudden and sometimes violent circumstances.

So it is easy to understand the reluctance of individual scientists, institutions and funding agencies to press too vocally for access to more brains from newborns, infants and older children, as well as for fetal brains obtained after abortion. Still, as a News Feature on page 442 shows, some patient groups in the United States are raising their voices. These groups deserve support. Moreover, they need scientific organizations in the United States and abroad to endorse and work towards a more ambitious goal: an international tissue bank holding perhaps tens of thousands of brains from young children and human fetuses around the world. *Nature* today pledges its support for such a bank.

The case for such a facility comes down to the growing number of scientists who wish to study brains from these early stages of development. Biological technologies now allow the extraction of a wealth of information about neurological diseases caused by faulty brain development. Neurodevelopmental conditions such as autism, schizophrenia and bipolar disorder are a huge societal burden, yet there are few effective treatments for them. Schizophrenia alone costs the United States tens of billions of dollars each year.

Many scientists working on these diseases have no access to young brain tissue. An informal survey of existing brain banks shows that they hold tissue from barely 1,300 brains collected from fetuses, young children and teenagers.

How can the supply to science of such sensitive material — the very seat of a child's personality — be increased on an international scale? The supply shortage is partly down to logistical problems. But these have been solved by some adult brain banks and, in principle, the logistics are no different when it comes to children.

The largest obstacle — real or perceived — remains the sensitivity of the subject and the difficulty of raising it with parents. Ultimately, this must be confronted, and discussions in the United States are being led by autism advocacy groups, who are used to talking to distressed parents. They are working to convince parents of the value of donating a child's brain, should the child die in an accident. In parallel, the US National Institutes of Health (NIH) has signed up in principle to establish a nationwide network of brain collection that will actively include those of children, and serve wider biomedical communities.

Things are moving slowly, but at least they are moving. And it is not too early to think how this slow domestic progress could be geared up to an international scale. More countries means more donors, and the

high natural variability between brains means that very large numbers can be needed to give studies of brain material statistical significance.

International networking of brain banks has been done before. BrainNet Europe, for example, was established 10 years ago as a single portal for brain tissue collected in 19 different European countries, none of which holds significant amounts of children's brains. The US advocacy organization Autism Speaks has already added a UK collection point in Oxford to its own brain bank.

***"It is understandable that bodies such as the NIH don't make strong public statements about extending fetal collections."***

Networked collections must also extend to fetal brains, and this is another reason for an international approach. At least eight people, including four doctors, associated with abortion procedures have been murdered for their activities in the United States in the past 20 years, so

it is understandable that bodies such as the NIH don't make strong public statements about extending fetal collections.

The Lieber Institute for Brain Development in Baltimore, Maryland, dedicated to translational research in neurodevelopmental diseases, which opened formally this year, is building a collection of young brains to support its research programme. It will begin its fetal acquisitions at three locations in Europe — Scotland, Denmark and Bulgaria — where the political climate is less hostile. In fact, the Lieber institute is proving a model for how to move ahead on paediatric and fetal brain banking, with appropriate respect for personal and political sensitivities. Would it cooperate with a national or international brain bank network, if created? Yes, say its organizers. But it won't wait for it. There is too much important research that needs to be done. ■

## Animal talk

*Germany must do more to encourage dialogue on animal experimentation.*

Sandwiched between the towering edifices of the Bundestag and the Chancellor's office, and just a short walk from other government buildings, the old family villa that is home to the Swiss embassy makes for a curious sight in the political heart of modern-day Berlin. Last week, the embassy hosted an international meeting of scientists from around the world who defend the use of animals in research. But despite being invited, nobody from the government offices bothered even to drop in. German animal-welfare groups also

declined to attend. That was unfortunate given that the gathering was intended to discuss the principles of the Basel Declaration, which promotes outreach by animal researchers to politicians and the public. And something else failed to materialize — Germany's plans to create a professional office to promote and implement the Basel Declaration principles, which some attendees had hoped would be announced at the meeting by the country's research organizations.

This lack of action reflects poorly on Germany's proclaimed interest in creating an environment within which its generously funded biomedical research can flourish. And it is disconcerting, because, like all countries in the European Union (EU), Germany must translate into national law a complex and controversial directive that regulates the use of animals in research.

The Basel Declaration was drafted at a meeting of mostly Swiss and German scientists last November. It has now been signed by nearly 900 people, some 500 of whom came from other countries. The scientists want the declaration to have the same authority over the ethics of animal experimentation as the 1964 Declaration of Helsinki has over the ethics of human experimentation. The formal infrastructure being developed around the declaration could help to realize this ambition.

The declaration was prompted by concerns over the EU animal-research directive, early drafts of which were so unfriendly to researchers that European scientists were shocked at how unprepared they were to lobby in the same arena as animal welfare and rights groups. Political battles raged for years before the directive was finally approved in 2010. Only one country abstained from what was otherwise a unanimous vote: Germany.

Why? Germany handed prime responsibility for the directive to the agricultural ministry. Others in the government, notably the research ministry, disagreed with this approach and the two ministries could

not agree on much right up to the vote. The agricultural ministry is now handling implementation without consultation with scientists.

Had representatives of the German government showed up at last week's meeting, as their Swiss counterparts did, they would have heard from researchers how the loose wording of much of the directive could create difficulties for them while it is being implemented, and how they must therefore be consulted.

**“Germany must translate into national law a complex and controversial directive.”**

For example, the directive requires that a ‘severity degree’ classification be introduced for all approved animal procedures. The signatories to the Basel Declaration approve of this, but some government offices in Europe have discussed whether an experiment should automatically be given a higher severity grade if it uses animals that have been genetically engineered, and whether the classifications should be made public.

German animal-welfare groups could also be part of these debates — as Swiss ones are — but they rarely communicate with the research community.

This relationship presents a challenge for German signatories to the Basel agreement and is a prime example of why Germany needs an office to coordinate the outreach the declaration calls for. The country's research funding organizations — particularly stalwarts such as the Max Planck Society, the German Research Foundation and the Helmholtz Society — need to move swiftly to create such an office.

Switzerland has dodged bullets aimed at its sturdy scientific base by animal-rights campaigners and opponents of genetic engineering in recent years, partly by maintaining excellent communication and transparency. Germany will find it even harder to bring these groups together — but even the longest journey must start with a short walk. ■

## Scientific climate

*Results confirming climate change are welcome, even when released before peer review.*

**G**lobal warming is really happening — really. There was no conspiracy or cover-up. Peer review did not fail and the scientists who have spent decades working out the best way to handle and process data turned out to know how to handle and process data after all. Thank you Berkeley Earth Surface Temperature (BEST) study.

Four papers released by the BEST team at the University of California, Berkeley, last week are of undoubted interest to the media, given that they support what is portrayed as the mainstream scientific position on climate change. They could also find traction in politics, especially in the United States, where they could be used to combat the assertions of Republicans, who have effectively tossed climate science away. But the headline scientific conclusion, that a century and a half of instrumental measurements confirm a warming trend, is, well, all a little 1990.

Of course, reproduction of existing results is a valid contribution, and the statistical methods developed by the BEST team could be useful additions to climate science. But valid contributions and useful additions alone do not generate worldwide headlines, so the massive publicity associated with the release of the papers (which were simultaneously submitted to the *Journal of Geophysical Research*) is a curious affair.

There was predictable grumbling at the media coverage from within the scientific community, which saw it as publicity in lieu of peer review. Reporters are more than happy to cover the story now, while it's sexy, but will they cover it later, when the results are confirmed, adjusted or corrected in accordance with a thorough vetting? The

short answer is no, many of them will not. Barring an extraordinary reversal of message, the wave of press coverage is likely to be only a ripple when the papers are finally published. And this is what upsets the purists: the communication of science in this case comes before the scientific process has run its course.

Members of the Berkeley team revelled in their role as scientific renegades. Richard Muller, the physicist in charge, even told the BBC: “That is the way I practised science for decades; it was the way everyone practised it until some magazines — particularly *Science* and *Nature* — forbade it.”

This is both wrong and unhelpful. It is wrong because for years *Nature* has explicitly endorsed the use of preprint servers and conferences as important avenues for scientific discussion ahead of submission to this journal, or other *Nature* titles. For example, on page 493 this week we publish a paper that discusses the dwarf planet Eris, based on results that the lead author presented (with *Nature's* knowledge and consent) at a conference several weeks ago. Journalists are, of course, welcome to report what they come across in such venues — as several did on Eris. What *Nature* discourages is authors specifically promoting their work to the media before a peer-reviewed paper is available for others in the field to read and evaluate.

Muller's statement is unhelpful because such inflammatory claims can only fuel the heated but misguided debate on climate-sceptic blogs and elsewhere about the way science works and how it treats those who insist on viewing themselves as outsiders.

To solicit input on results before publication is not a guerrilla action against a shadowy scientific elite. Witness the posting on a preprint server last month of the paper reporting neutrinos that apparently travel faster than light: the authors made it clear that they were seeking

help from the wider community to explain the findings, and the media stories (if not the headlines) mostly reflected that. To pretend otherwise can only erode public trust in science, as it is practised by all. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunjv](http://go.nature.com/xhunjv)



## Piecemeal cuts won't add up to radical reductions

*To meet ambitious emissions targets will require systems thinking and massive breakthroughs in technology and fuels, says Jane C. S. Long.*

In 2005, California threw down the gauntlet: by executive order, the state must reduce greenhouse-gas emissions to 80% below what they were in 1990, by 2050. Similar targets have been adopted in Europe, but the California goal is well beyond any federal policy taken on in the United States. Is it possible? What will it take to achieve it? For two years, I was part of a group of energy experts in California that tried to answer those questions. Our report, *California's Energy Future — The View to 2050*, was released by the California Council on Science and Technology earlier this year.

For smaller emissions-reduction targets, tactical approaches such as piecemeal reductions may look promising. But to ensure a radical decrease in emissions while also reliably meeting its energy needs, California must make strategic choices.

The difference may seem academic, but in fact it is hugely significant. For example, if your net emissions target is not near-zero, you might approach it by increasing the use of biofuels in cars. But biofuels are scarce. To achieve near-zero net emissions, you must electrify the cars so that you can reserve the biofuels for forms of transport that cannot be electrified — heavy-duty trucks and planes, for example. In a systems approach, using biofuels in cars is a dead end.

Similarly, it is popular to promote extensive use of wind power, with no worries about what to do when the wind doesn't blow. Somehow the problem just gets 'solved'. Some say that we can radically reduce emissions with only a major emphasis on efficiency, or just by changing our behaviour. But what if it doesn't add up?

In our report, the California's Energy Future Committee looked at the big picture, asking which technical strategies will achieve an energy system with near-zero emissions yet still meet society's needs. We estimated how much more efficient buildings, industry and transportation could become, and how quickly cars, buses, trains and heat production could be electrified. We looked at how to supply that electricity from near-zero-emissions sources: nuclear power, fuel-based power plants used with carbon capture and storage technology, and renewable energy. We also worried about emissions from 'load balancing', in which generators are used to meet peak loads or fill in for intermittent power from solar or wind sources. We assessed how much biomass might be sustainably available to meet the remaining demand for fuel, and how much it could help to cut emissions. We counted everything, but only once. It was hard, but it was honest.

Having done the maths, what did we discover?

If California could very quickly replace cars, appliances, boilers, buildings and power plants with today's state-of-the-art technology, replace and expand current electricity generation with

non-emitting sources and produce as much biofuel as possible by 2050, the state could reduce emissions a lot — by perhaps 60% below 1990 levels. But it would have to replace or retrofit every building to very high efficiency standards. Electricity would have to replace natural gas for home and commercial heating. All buses and trains, virtually all cars, and some trucks would be electric or hybrid. And the state's entire electricity-generation capacity would have to be doubled, while simultaneously being replaced with emissions-free generation. Low-emissions fuels would have to be made from California's waste biomass plus some fuel crops grown on marginal lands without irrigation or fertilizer.

To reach an 80% cut will take new technology.

Emissions-free electricity is one hurdle. California has plenty of renewable resources, but they are intermittent. Energy-storage technology is not yet good enough to solve this problem, and no one knows

whether smart-grid technologies can. Using natural-gas generators to firm up the supply will mean falling short of the 80% goal.

A reliable reinvented energy system should provide base-load power without intermittency or emissions. California should exploit all the geothermal energy it can. Carbon-capture schemes should focus not on coal-fired plants, but on lower-cost natural-gas plants, which produce fewer emissions to sequester. And the state should rethink its opposition to nuclear power.

Even if the electricity problem can be solved, it won't address the needs of planes, trucks, ships and some industrial heating that cannot be electrified. The state will still need fuel — about three-quarters as much as today. California

would be lucky to get half of that from biofuels.

So there we are — a concerted effort to deploy known technology could cut emissions by more than half, but getting all the way down to 80% cuts will almost certainly require major advances in near-zero-emissions fuels. This is by far the biggest technology gap. The conclusion may seem obvious, but few have really given this the hardheaded look it deserves. California can't just spend or deploy its way to an 80% reduction or beyond — and neither can anywhere else.

We don't know precisely how economic and political factors will help or hinder progress towards the reduction target. But we are obliged to try to reach it, and we now know what it will take. This is not a small thing. We may not make the goal of radical emissions cuts by 2050, but it is important to get there eventually — or rather, as fast as we can. ■

**Jane C. S. Long** is principal associate director at large at the Lawrence Livermore National Laboratory in Livermore, California, USA.  
e-mail: [janeclong@gmail.com](mailto:janeclong@gmail.com)

GETTING ALL THE WAY  
DOWN TO  
**80% CUTS**  
WILL ALMOST  
CERTAINLY REQUIRE  
**MAJOR  
ADVANCES.**

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/rv9yuu](http://go.nature.com/rv9yuu)

# SEVEN DAYS

The news in brief

## POLICY

### Science Europe

A new Brussels-based lobby group, Science Europe, held its founding assembly in Berlin on 21 October. It hopes to become the “single voice for science in Europe”, its president, Paul Boyle, told *Nature* in August (see *Nature* 477, 18; 2011). The organization ([www.scienceeurope.org](http://www.scienceeurope.org)) unites two science-advocacy groups: the European Heads of Research Councils, which has now been officially dissolved, and the European Science Foundation, which is continuing as a separate body but may wind down its activities over the next few years.

### No carbon capture

The United Kingdom's energy and climate-change agency has ditched plans to invest £1 billion (US\$1.6 billion) in a project to capture and bury millions of tonnes of carbon dioxide produced each year by a coal-fired power station in Longannet, Scotland. The project, which aimed to be even larger than the US flagship FutureGen programme, was once a front runner in Britain's much-delayed competition to receive funding for carbon-capture and storage schemes. The money will be spent on other carbon-capture projects, the agency said on 19 October. See [go.nature.com/xqncd](http://go.nature.com/xqncd) for more.

## RESEARCH

### A planet is born

Astronomers have for the first time imaged a planet so young — a mere 2 million years old — that it is still gathering material from its birth site, a disk of gas and dust surrounding a star 145 parsecs (473 light years) from Earth. It lies about as far from its star

as Uranus does from our Sun. The finding was published on the arXiv preprint server on 17 October (A. L. Kraus and M. J. Ireland preprint at <http://arxiv.org/abs/1110.3808v1>; 2011), although it had been shared at conferences previously. See [go.nature.com/opa1pd](http://go.nature.com/opa1pd) for more.

### Warming verified

An independent analysis of the land-surface-temperature record has concluded — if anyone was in doubt — that global warming is happening. The Berkeley Earth Surface Temperature (BEST) study, led by Richard Muller, a physicist at the University of California, Berkeley, mostly agreed with results from three teams that had previously studied

data from land-temperature stations, although it used different statistical methods for the analysis. The BEST analysis was released on 20 October, but no part of it has yet been peer reviewed; the team is preparing to submit four papers to the *Journal of Geophysical Research*. See [go.nature.com/kqiypu](http://go.nature.com/kqiypu) and page 428 for more.

### Malaria vaccine

The world's leading candidate for a malaria vaccine has claimed promising results in a phase III trial, from which the first findings were published on 18 October. But the RTS,S/AS01 vaccine, funded by pharmaceutical company GlaxoSmithKline and the global PATH Malaria Vaccine Initiative, showed low

the Arabian tectonic plate converges with Eurasia, is prone to devastating earthquakes: a magnitude-7.6 quake in 1999 killed 17,000 people; and in 1976, a magnitude-7.3 quake struck just 70 kilometres away from this week's epicentre and killed several thousand.



A. ANTAKYALU/AP

## Earthquake strikes eastern Turkey

Several hundred people are feared to have died in the magnitude-7.2 earthquake that hit eastern Turkey on 23 October. As *Nature* went to press, rescue teams were searching for survivors in cities such as Erciş and Van, near the Iranian border. The region, where

efficacy against severe forms of malaria, disappointing some experts. See page 439 for more.

### Hope for MS drug

A monoclonal-antibody treatment for multiple sclerosis has seen positive results in a late-stage clinical trial. The results, presented on 22 October at a congress in Amsterdam, found that 78% of patients treated with alemtuzumab remained free from relapse (a flare-up of inflammation) after two years, compared with 59% using one of the standard therapies, interferon  $\beta$ -1a. But evidence that the drug can actually reverse nerve damage was not statistically significant (unlike in earlier trials). Alemtuzumab is made

by Genzyme, a US subsidiary of Paris-based drug maker Sanofi. See [go.nature.com/suucta](http://go.nature.com/suucta) for more.

## BUSINESS

## Abbott splits up

Abbott Laboratories is splitting into two companies: an as-yet-unnamed research-based pharmaceuticals firm, and a medical-products business covering everything from generic drugs to lab diagnostics. The drug-maker, based in Chicago, Illinois, is the world's ninth-largest in terms of global revenues, bringing in US\$35.2 billion in 2010. The separation, announced on 19 October, is widely seen as an attempt to attract health-care investors to a business free of branded drugs, products on which some analysts feel the firm is overly dependent.

## EVENTS

## Floods in Thailand

Northern suburbs in Thailand's capital Bangkok were last week inundated by heavy floods (pictured), described as the country's worst in half a century. As *Nature* went to press, officials were hoping that floodwaters would retreat, and that defensive walls and drainage canals would save the centre of the low-lying city from severe



damage. The floods, which have been going on since late July, have killed more than 350 people and caused billions of dollars in damage.

## Europe GPS launch

The European Space Agency has launched the first two operational spacecraft of Galileo, Europe's global positioning system. The satellites, launched on 21 October, joined two test satellites already in orbit. The network, costing more than €5 billion (US\$6.9 billion), will feature up to 27 operational satellites and three spares. See [go.nature.com/jvcdb1](http://go.nature.com/jvcdb1) for more.

## PEOPLE

## Stealing secrets

A scientist who was born in China but is a permanent resident in the United States has pleaded guilty to economic espionage and stealing trade

secrets from two former US employers to benefit Hunan Normal University in China. Kexue Huang admitted to passing on details about pesticides that he had learned when working at Dow AgroSciences in Indianapolis, Indiana, from 2003 to 2008. He also admitted to stealing a key component of a food product developed at his subsequent employer, grain distributor Cargill in Minneapolis, Minnesota, federal prosecutors said on 18 October.

## Journal chief

Molecular biologist Inder Verma will be the next editor-in-chief of the *Proceedings of the National Academy of Sciences*, the journal announced on 19 October. Verma, currently at the Salk Institute in La Jolla, California, replaces Randy Schekman, editor since 2006, who is moving to edit a new open-access life-sciences journal (see *Nature* 475, 145; 2011).

## Head for NIH centre

Cell biologist Chris Kaiser will be the next director of the basic-biosciences institute at the US National Institutes of Health in Bethesda, Maryland, the agency announced on 18 October. Kaiser, who currently heads the biology department at the Massachusetts Institute of Technology in Cambridge, will

## COMING UP

### 31 OCTOBER

The United Nations proclaims that the world population has reached 7 billion — see *Nature* 478, 300 (2011) for more.

### 1 NOVEMBER

China's Shenzhou 8 spacecraft is rumoured to launch. It is the country's first attempt to remotely dock a craft with its Tiangong 1 space module, which launched last month.

### 2–4 NOVEMBER

Progress in using stem cells to regrow livers, lungs, kidneys and spinal cords is discussed at the World Conference on Regenerative Medicine in Leipzig, Germany. [go.nature.com/jsdbdy](http://go.nature.com/jsdbdy)

lead the US\$2-billion National Institute of General Medical Sciences from spring 2012. He replaces acting director Judith Greenberg, who has been covering the role since former director Jeremy Berg stepped down in July. See [go.nature.com/a2zmbi](http://go.nature.com/a2zmbi) for an interview with Kaiser.

## UK physics head

Britain's most financially troubled science funding agency, the Science and Technology Facilities Council, has a new chief executive: John Womersley, who was its director of science programmes. He replaces Keith Mason, who was criticized for poor community engagement during years of dire financial straits for the council, which was founded by the merger of two councils in 2007 and funds mainly physics and astronomy research (see *Nature* 462, 396; 2009). The appointment was announced on 18 October.

➔ [NATURE.COM](http://NATURE.COM)

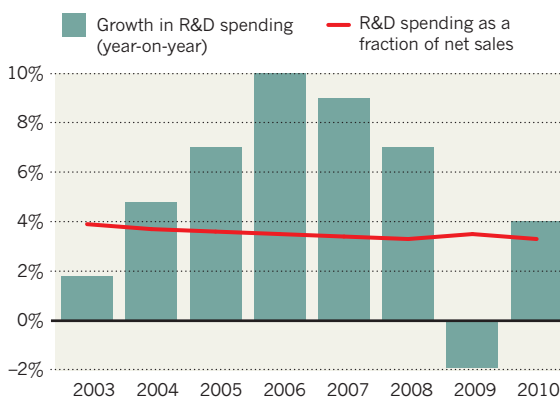
For daily news updates see: [www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

According to the European Commission's 2011 scoreboard for industry spending on research and development (R&D), released on 18 October, the world's top 1,400 companies increased their R&D investment by 4% last year, after a 1.9% drop in 2009. Swiss drug firm Roche, based in Basel, spent €7.2 billion (US\$10 billion) on research — the most of any firm. The rise broadly mirrors a similar upturn in sales. Research spending as a proportion of net sales has fallen slowly over the past half decade (see chart).

## CORPORATE RESEARCH SPENDING RECOVERS

Firms increased their spending on research and development (R&D) last year, although the recovery is largely in line with greater sales.



# NEWS IN FOCUS

**ARSENIC** Scientists seek the source of contaminated groundwater **p.437**

**GLOBAL HEALTH** Researchers raise questions over malaria vaccine **p.439**

**STEM CELLS** The bright side of Europe's lamented patent ruling **p.441**

**NEUROSCIENCE** The shortage of children's brains in research **p.442**



AP PHOTO/S. PONOMAREV

The Fukushima accident led to mass evacuations from nearby towns such as Minamisoma.

## NUCLEAR DISASTER

# Fallout forensics hike radiation toll

*Global data on Fukushima challenge Japanese estimates.*

BY GEOFF BRUMFIEL

The disaster at the Fukushima Daiichi nuclear plant in March released far more radiation than the Japanese government has claimed. So concludes a study<sup>1</sup> that combines radioactivity data from across the globe to estimate the scale and fate of emissions from the shattered plant.

The study also suggests that, contrary to government claims, pools used to store spent nuclear fuel played a significant part in the release of the long-lived environmental contaminant caesium-137, which could have been prevented by prompt action. The analysis has been posted online for open peer review by the journal *Atmospheric Chemistry and Physics*.

Andreas Stohl, an atmospheric scientist with the Norwegian Institute for Air Research in Kjeller, who led the research, believes that

the analysis is the most comprehensive effort yet to understand how much radiation was released from Fukushima Daiichi. "It's a very valuable contribution," says Lars-Erik De Geer, an atmospheric modeller with the Swedish Defense Research Agency in Stockholm, who was not involved with the study.

The reconstruction relies on data from dozens of radiation monitoring stations in Japan and around the world. Many are part of a global network to watch for tests of nuclear weapons that is run by the Comprehensive Nuclear-Test-Ban Treaty Organization in Vienna. The scientists added data from independent stations in Canada, Japan and Europe, and then combined those with large European and American caches of global meteorological data.

Stohl cautions that the resulting model is far from perfect. Measurements were scarce in the immediate aftermath of the Fukushima

accident, and some monitoring posts were too contaminated by radioactivity to provide reliable data. More importantly, exactly what happened inside the reactors — a crucial part of understanding what they emitted — remains a mystery that may never be solved. "If you look at the estimates for Chernobyl, you still have a large uncertainty 25 years later," says Stohl.

Nevertheless, the study provides a sweeping view of the accident. "They really took a global view and used all the data available," says De Geer.

## CHALLENGING NUMBERS

Japanese investigators had already developed a detailed timeline of events following the 11 March earthquake that precipitated the disaster. Hours after the quake rocked the six reactors at Fukushima Daiichi, the tsunami arrived, knocking out crucial diesel back-up generators designed to cool the reactors in an emergency. Within days, the three reactors operating at the time of the accident overheated and released hydrogen gas, leading to massive explosions. Radioactive fuel recently removed from a fourth reactor was being held in a storage pool at the time of the quake, and on 14 March the pool overheated, possibly sparking fires in the building over the next few days.

But accounting for the radiation that came from the plants has proved much harder than reconstructing this chain of events. The latest report from the Japanese government, published in June, says that the plant released  $1.5 \times 10^{16}$  becquerels of caesium-137, an isotope with a 30-year half-life that is responsible for most of the long-term contamination from the plant<sup>2</sup>. A far larger amount of xenon-133,  $1.1 \times 10^{19}$  Bq, was released, according to official government estimates.

The new study challenges those numbers. On the basis of its reconstructions, the team claims that the accident released around  $1.7 \times 10^{19}$  Bq of xenon-133, greater than the estimated total radioactive release of  $1.4 \times 10^{19}$  Bq from Chernobyl.

The fact that three reactors exploded in the Fukushima accident accounts for the huge xenon tally, says De Geer.

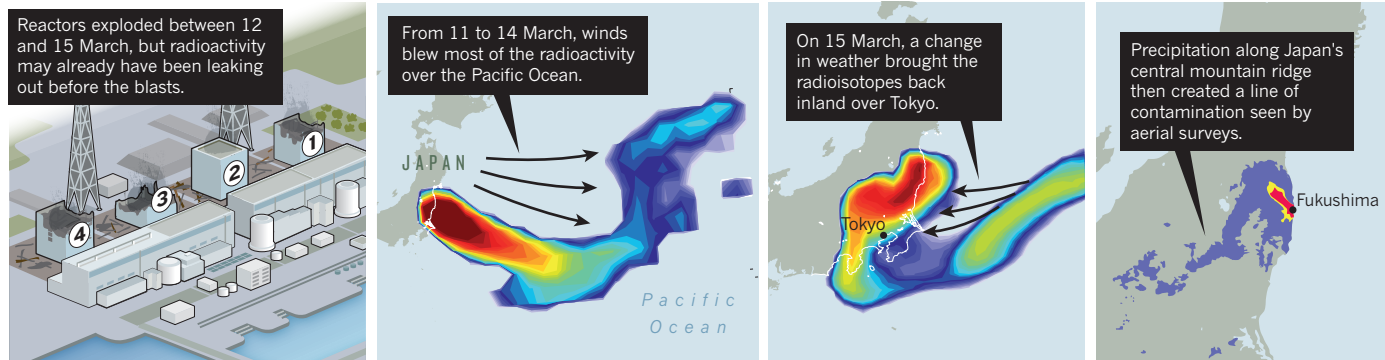
Xenon-133 does not pose serious health risks because it is not absorbed by the body ▶



**FUKUSHIMA CRISIS**  
➔ [WWW.NATURE.COM/JAPANQUAKE](http://WWW.NATURE.COM/JAPANQUAKE)

## RADIOISOTOPE RECONSTRUCTION

After a massive earthquake and tsunami hit Japan on 11 March, three reactors at Fukushima Daiichi blew up and a fourth caught fire. A reconstruction now shows how radioisotopes streamed from the plant and swept across the country.



SOURCE: REF. 1/MEXT

► or the environment. Caesium-137 fallout, however, is a much greater concern because it will linger in the environment for decades. The new model shows that Fukushima released  $3.5 \times 10^{16}$  Bq caesium-137, roughly twice the official government figure, and half the release from Chernobyl. The higher number is obviously worrying, says De Geer, although ongoing ground surveys are the only way to truly establish the public-health risk.

Stohl believes that the discrepancy between the team's results and those of the Japanese government can be partly explained by the larger data set used. Japanese estimates rely primarily on data from monitoring posts inside Japan<sup>3</sup>, which never recorded the large quantities of radioactivity that blew out over the Pacific Ocean, and eventually reached North America and Europe. "Taking account of the radiation that has drifted out to the Pacific is essential for getting a real picture of the size and character of the accident," says Tomoya Yamauchi, a radiation physicist at Kobe University who has been measuring radioisotope contamination in soil around Fukushima.

Stohl adds that he is sympathetic to the Japanese teams responsible for the official estimate. "They wanted to get something out quickly," he says. The differences between the two studies may seem large, notes Yukio Hayakawa, a volcanologist at Gunma University who has also modelled the accident, but uncertainties in the models mean that the estimates are actually quite similar.

The new analysis also claims that the spent fuel being stored in the unit 4 pool emitted copious quantities of caesium-137. Japanese officials have maintained that virtually no radioactivity

leaked from the pool. Yet Stohl's model clearly shows that dousing the pool with water caused the plant's caesium-137 emissions to drop markedly (see 'Radiation crisis'). The finding implies that much of the fallout could have been prevented by flooding the pool earlier.

The Japanese authorities continue to maintain that the spent fuel was not a significant source of contamination, because the pool itself did not seem to suffer major damage. "I think the release from unit 4 is not important," says Masamichi Chino, a scientist with the Japanese Atomic Energy Authority in Ibaraki, who helped to develop the Japanese official estimate. But De Geer says the new analysis implicating the fuel pool "looks convincing".

The latest analysis also presents evidence that xenon-133 began to vent from Fukushima Daiichi immediately after the quake, and before the tsunami swamped the area.

This implies that even without the devastating flood, the earthquake alone was sufficient to cause damage at the plant.

The Japanese government's report has already acknowledged that the shaking at Fukushima Daiichi exceeded the plant's design specifications. Anti-nuclear activists have long been concerned that the government has failed to adequately address geological hazards when licensing nuclear plants (see *Nature* **448**, 392–393; 2007), and the whiff of xenon could prompt a major rethink of reactor safety assessments, says Yamauchi.

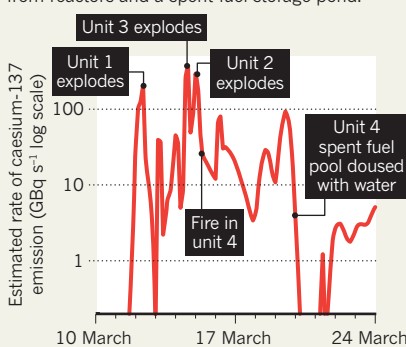
The model also shows that the accident could easily have had a much more devastating impact on the people of Tokyo. In the first days after the accident the wind was blowing out to sea, but on the afternoon of 14 March it turned back towards shore, bringing clouds of radioactive caesium-137 over a huge swathe of the country (see 'Radioisotope reconstruction'). Where precipitation fell, along the country's central mountain ranges and to the northwest of the plant, higher levels of radioactivity were later recorded in the soil; thankfully, the capital and other densely populated areas had dry weather. "There was a period when quite a high concentration went over Tokyo, but it didn't rain," says Stohl. "It could have been much worse." ■

Additional reporting by David Cyranoski and Rina Nozawa.

1. Stohl, A. *et al.* *Atmos. Chem. Phys. Discuss.* **11**, 28319–28394 (2011).
2. [www.kantei.go.jp/foreign/kan/topics/201106/iaea\\_houkokusho\\_e.html](http://www.kantei.go.jp/foreign/kan/topics/201106/iaea_houkokusho_e.html)
3. Chino, M. *et al.* *J. Nucl. Sci. Technol.* **48**, 1129–1134 (2011).

## RADIATION CRISIS

Modelling the first week of the Fukushima disaster reveals that huge bursts of radioisotopes poured from reactors and a spent-fuel storage pond.



SOURCE: REF. 1

  
**MORE  
ONLINE**

## TOP STORY

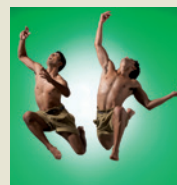


Brazilian scientists fight for cut of oil royalties  
[go.nature.com/r5bqva](http://go.nature.com/r5bqva)

## OTHER NEWS

- Finding puts brakes on faster-than-light neutrinos [go.nature.com/sitdhx](http://go.nature.com/sitdhx)
- Mastodon fossil throws up questions over 'rapid' extinction [go.nature.com/1a71te](http://go.nature.com/1a71te)
- California finalizes climate regulations [go.nature.com/7bk8hb](http://go.nature.com/7bk8hb)

## Q&amp;A



A leading dancer and a Cambridge scientist discuss their collaboration  
[go.nature.com/scnp1t](http://go.nature.com/scnp1t)

H. GLENDINNING

## ENVIRONMENTAL SCIENCE

# Conflicting studies fuel arsenic debate

*Uncertainties over contaminated groundwater in southern Asia highlight gaps in science.*

BY EUGENIE SAMUEL REICH

It is often called the largest mass poisoning in human history. Yet more than a decade after the discovery that drinking-water wells in Bangladesh are contaminated with naturally occurring arsenic, researchers are still struggling to understand the hydrology and chemistry behind an environmental crisis that affects more than 60 million people in the Bengal Basin.

A paper<sup>1</sup> in the press in *Geophysical Research Letters* adds a further piece to the puzzle. It concludes that one factor, man-made ponds, probably does not play a major part in the release of arsenic into groundwater as some geologists have claimed. Yet the mechanism that is responsible remains elusive, in part because of a lack of systematic studies of the region affected by the problem.

Groundwater in sites across south and east Asia contains arsenic at levels that can cause chronic poisoning. Symptoms include skin lesions, respiratory and cardiovascular disease and a range of cancers. From the basins of the Indus River in the west to the Yangtze in China, “almost all the major rivers draining the Himalayas are affected”, says Scott Fendorf of Stanford University in California, a biogeochemist who has studied the problem in Cambodia. Residual tectonic activity from the formation of the Himalayas and high rainfall in the region combine to produce rapid erosion, which sweeps arsenic — associated with iron pyrite ores — along with iron oxide particles into river sediment below.

The public-health consequences are most stark in the Ganges delta of Bangladesh, where, in the 1970s, aid agencies encouraged the drilling of hand-pumped wells so that people did not have to drink surface water contaminated with waterborne microbial diseases such as cholera. Geologists failed to realize that the pumps would be tapping into arsenic-laced aquifers beneath.

But just how the arsenic gets into the groundwater is not known. In 2008, a group led by hydrologist Charles Harvey at the Massachusetts Institute of Technology in Cambridge reported<sup>2</sup> that water from an aquifer in the Munshiganj district of Bangladesh has the same isotopic fingerprint as pond water there, suggesting that organic material may be seeping



REUTERS/A BRAJ

High levels of arsenic in drinking water threaten the health of millions of people living in the Bengal Basin.

from ponds into the aquifer 30 metres below. Harvey proposed that the arsenic is released into the water supply when organic material from ponds triggers microbial activity, which can dissolve the iron oxide particles and release their load of arsenic into the water supply.

The man-made ponds in Bangladesh, a by-product of excavating earth to build up the land against floodwater caused by the monsoon rains, cover a 22-fold greater area than they did 50 years ago, making them a plausible suspect. Harvey and his colleagues suggest that contamination might be reduced if ponds were not dug near wells and if wells were not drilled near existing ponds.

But the finding proved controversial after a group led by geochemist John McArthur at University College London — who failed to see the effect while studying arsenic contamination in West Bengal, India, in 2004 — criticized some of Harvey's team's methods<sup>3</sup>. McArthur and his colleagues reanalysed Harvey's data and arrived at the opposite conclusion. They argued that arsenic was being produced in sediments near

the aquifer itself by long-buried organic matter, and that water from ponds was actually flushing out the aquifer and reducing contamination.

The latest study<sup>1</sup> also suggests that the ponds are not to blame. In their paper, Karen Johannesson of Tulane University in New Orleans, Louisiana, and her colleagues report that six arsenic-laden sites west of the Ganges in West Bengal fail to show signs of organic matter from ponds. The researchers also argue that organic matter from ponds would take thousands of years to penetrate through the tens of metres of clay and sand to the aquifers below, far longer than Bangladeshis have been digging ponds. “I think we feel pretty confident, at least where we're working, that ponds are not contributing,” Johannesson says.

But Harvey's and Johannesson's groups worked at widely separated sites, where conditions may differ (see map), says hydrologist Roger Beckie of the University of British Columbia in Vancouver. He suspects that some organic matter is coming from ponds whereas some is intrinsic to the sediments. “My gut feeling is that both processes are at work,” he says.

Hydrologist Abhijit Mukherjee at the Indian Institute of Technology in Kharagpur, who is working with West Bengal's government to find ways of predicting where safe wells

**“Almost all the major rivers draining the Himalayas are affected.”**

► might be dug, says the dispute shows that government and funding agencies need to support studies at multiple sites, from the Himalayas to the Bay of Bengal. “If that is done we have a systematic way of comparing data from sites,” he says.

Whatever the origin of the arsenic, says Stephan Hug, a geochemist at the Swiss Federal Institute of Aquatic Science and Technology in Dübendorf, one possible solution would be to drill deep wells — to below around 150 metres — where the water should be arsenic free. But then if too many such wells are drilled, the arsenic could

### TOXIC WATERS

Arsenic transported from the Himalayas by the Ganges and other rivers has found its way into groundwater in the Bengal Basin, but three studies using data from three separate locations (inset) disagree on how.



be pulled down to that level, he adds. Moreover, water from such depths can be unacceptably saline.

Mukherjee says that the main goal should be to learn how the arsenic is getting into the groundwater, so that researchers can make wells safe by ensuring that such conditions are avoided. The drinking water of millions of people depends on it. ■

1. Datta, S. *et al. Geophys. Res. Lett.* <http://dx.doi.org/10.1029/2011GL049301> (2011).
2. Neumann, R. B. *et al. Nature Geosci.* **3**, 46–52 (2010).
3. McArthur, J. M., Ravenscroft, P. & Sracek, O. *Nature Geosci.* **4**, 655–656 (2011).

### REMUNERATION

# US bill targets grantee salaries

*Cost-saving measure would reduce maximum amount paid to biomedical researchers funded by federal agencies.*

BY MEREDITH WADMAN

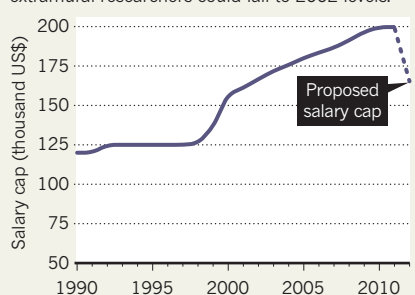
The US House of Representatives is considering legislation that would roll back the maximum amount of a grant that can go towards the salary of a biomedical researcher funded by a federal agency. Under a spending bill proposed on 29 September, the ‘salary cap’ would return to what it was roughly ten years ago. If Congress passes the measure, universities and other institutions will face a dilemma: dip into other funds to make up the difference and keep leading investigators happy, or risk losing researchers to industry, clinical practice or competing institutions.

The 2012 spending bill would cut the salary cap by 17%, from US\$199,700 to \$165,300, for extramural scientists funded by the National Institutes of Health (NIH; see ‘Lowering the ceiling’), the Centers for Disease Control and Prevention and other agencies in the Department of Health and Human Services (DHHS). A parallel Senate bill leaves the salary cap untouched; the two versions must be harmonized before a final 2012 DHHS spending bill is passed. The measures will be considered as part of fraught budget negotiations for financial year 2012.

Just how many scientists would be affected is not clear, but the cut would certainly hit many of the thousands of principal investigators who receive grants from the NIH.

### LOWERING THE CEILING

The salary cap for US National Institutes of Health extramural researchers could fall to 2002 levels.



Among medical schools and teaching hospitals, “there is very strong concern about this proposal”, says David Moore, senior director for governmental relations at the Association of American Medical Colleges (AAMC) in Washington DC. “If you are a research-intensive institution and you’re talking about hundreds of faculty where you have to make up this difference, we’re talking about millions of dollars the institution is now responsible for.”

On 11 October, the AAMC co-authored a letter to congressional-spending leaders including Denny Rehberg (Republican, Montana), who authored the House bill and is chairman of the subcommittee that funds the DHHS. Signed by 111 organizations and institutions, the letter argues that lowering the

salary cap would be most damaging to those investigators who dedicate the most time to research and therefore are most reliant on grants; would discourage gifted young scientists from entering research; and would drive physicians out of research even as demands for cures and therapies are growing.

“Physicians have alternative ways of surviving,” notes Moore. “If you continue to put more and more disincentives into a research career, at what point do those individuals say: ‘I’m going to go off and see patients?’”

Rex Chisholm, vice-dean for scientific affairs and graduate education at Northwestern University’s Feinberg School of Medicine in Chicago, Illinois, says that trying to make up for the cut “would create real difficulties for Northwestern and most other medical schools. Most of us don’t have that kind of money just lying around.”

He adds that in 2010, Feinberg, which is a signatory to the letter to Congress, paid nearly \$8.3 million towards the salaries of its 235 faculty members who earned more than the \$199,700 cap. Had the cap been \$165,300, he says, the medical school would have had to pay almost \$12 million to 283 faculty members.

Rehberg did not respond to *Nature’s* requests for comment.

If the cut goes through, says Jeff Gerber, a physician-researcher at the Children’s Hospital of Philadelphia in Pennsylvania, “you’re either going to have fewer [academic] positions available because the institutions will only be able to pay so many people, or, I suppose, money sometimes drives people into different professional choices”. Gerber says that his love for research will probably prevent him ever leaving it. But he expects hospitals to ask physicians to do less research and more income-generating clinical care if grant salaries are cut.

Ironically, the House bill that mandates the cut is far more generous to the NIH overall than is its Senate counterpart. The House bill would boost the biomedical agency’s budget by \$1 billion, or 3.3%, to \$31.7 billion. The Senate bill would cut it by \$190 million. ■

SOURCE: NIH



The RTS,S/AS01 candidate vaccine offers poor protection against severe malaria.

## GLOBAL HEALTH

# Malaria vaccine results face scrutiny

*Experts question early release of incomplete trial data.*

BY DECLAN BUTLER

“Malaria vaccine could save millions of children’s lives”; “World’s first malaria vaccine works in major trial”; “Malaria vaccine almost here”. To judge from last week’s headlines, scientists had made a big breakthrough in the long campaign to create a malaria vaccine, proving its effectiveness with interim results from a huge phase III clinical trial in Africa<sup>1</sup>.

Yet several leading vaccine researchers, who are critical of the unusual decision to publish partial trial data, argue that the results raise questions about whether the RTS,S/AS01 candidate vaccine can actually win approval.

RTS,S has been in development for some 25 years, initially by the US military, and since 2001 by a public-private venture between the PATH Malaria Vaccine Initiative (MVI) and the drug-maker GlaxoSmithKline (GSK), supported by US\$200 million in funding from the Bill & Melinda Gates Foundation. Bill Gates himself announced the interim results at the Gates Malaria Forum in Seattle, Washington.

Gates’ speech and the MVI’s public-relations material were suitably circumspect about the results, but they were “immediately translated into headlines about [reductions] in death and mortality”, says Andrew Farlow, an economist at the University of Oxford, UK, who has previously assessed the RTS,S programme<sup>2</sup>. “But the data are not telling you that at all.”

Some researchers question whether the results should have been published before all

the data were available; full results are expected in 2014. Interim trial data are usually reported only to regulatory authorities, and clinical trials published only once all the data are in, noted Nicholas White, a malaria expert at Mahidol University in Bangkok, in an editorial<sup>3</sup> accompanying the interim results. “There does not seem to be a clear scientific reason why this trial has been reported with less than half the efficacy results available,” he wrote.

The publication presents vaccine-efficacy data for infants aged 5–17 months, but not for those aged 6–12 weeks, who are the stated target of the trial: it is this group that would receive the malaria vaccine alongside routine immunizations. The aim of the trial is to provide the World Health Organization (WHO) with the information it needs to consider licensing the vaccine, and recommend it for use in that age group. “What is the point of publishing the interim data on the 5–17-month-olds?” asks Stephen Hoffman, a veteran malaria researcher and chief executive of a rival vaccine effort, Sanaria, based in Rockville, Maryland.

The MVI’s director, Christian Loucq, argues that the results were “robust enough to be published. We decided this before we knew the results; we felt it was our scientific and ethical duty to make the results public when they become available.”

One of the biggest claims made in the paper is that RTS,S reduced the total number of episodes of clinical malaria in the older group by 55.1%, ►

► NATURE.COM  
Vaccines special:  
[nature.com/vaccines](http://nature.com/vaccines)

► compared to controls. This measure of efficacy is recommended for assessing a partially effective vaccine<sup>4</sup>. But the public expects vaccine efficacy to describe protection over a period of time, argues Judith Epstein, a captain and paediatrician at the US Military Malaria Vaccine Program in Silver Spring, Maryland. Recalculating the trial data shows that RTS,S protected just 35–36% after 12 months, she says, adding that the paper should have presented both numbers. The study also showed no detectable impact on mortality, and it is too early to tell whether RTS,S actually protects against malaria, or merely delays infection.

The paper did report that RTS,S reduced severe malaria by 47% in the older group. But combining that result with available data from the younger age group cut that number to 34.8% — meaning that for the youngest children, the benefit must be even smaller. “The real question mark is the 34.8% efficacy in severe disease,” says Blaise Genton of the Swiss Tropical and Public Health Institute in Basel, and a member of the WHO technical advisory group for RTS,S. The results suggest that the vaccine might fall short of expectations, laid out in 2006 by a WHO-led consortium<sup>5</sup>, that it should have a “protective efficacy of more than 50% against severe disease and death and lasts longer than one year”. “If it doesn’t reduce deaths, and has only a modest effect on severe malaria, these are going to be big questions for decision-makers at WHO, GSK and the Gates Foundation,” says Hoffman.

Another worrying finding is that the frequency of serious adverse events, such as convulsions and meningitis, was significantly higher in the vaccinated group, although the data are too preliminary to draw firm conclusions. “The severe disease findings are a concern,” says Genton.

But Hoffman, like many researchers contacted by *Nature*, says that RTS,S still marks a significant achievement. It is the first vaccine against a complex multicellular parasite, *Plasmodium falciparum*, to consistently show a significant protective effect in large-scale trials. The phase III trial of RTS,S resulted in groundbreaking cooperation with African scientists, who led the 11 trials in 7 countries, says Hoffman. “I think that those teams deserve an incredible amount of recognition and congratulation.” ■

1. The RTS,S Clinical Trials Partnership *New Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1102287> (2011).

2. Farlow, A. In *The Science, Economics, and Politics of Malaria Vaccine Policy* (2006). Chapter available at <http://go.nature.com/8oxle3>

3. White, N. J. *New Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMe1111777> (2011).

4. Moorthy, V. S., Reed, Z. & Smith, P. G. *Vaccine* **27**, 624–628 (2009).

5. World Health Organization et al. *Malaria Vaccine Technology Roadmap* (2006). Available at <http://go.nature.com/iljdf7>

## GENETICS

# Fetal gene screening comes to market

*Non-invasive procedure could make prenatal testing easier, but it comes with ethical problems.*

BY ERIKA CHECK HAYDEN

Until last week, scrutinizing a fetus’s DNA for indications of genetic abnormalities meant tapping into the mother’s womb with a needle. Now there’s a test that can do it using a small sample of the mother’s blood. MaterniT21, a Down’s syndrome test that Sequenom of San Diego, California, launched in major centres across the United States on 17 October, is the first of several such tests expected on the market in the next year. It signals the arrival of a long-anticipated era of non-invasive prenatal genetic screening, with its attendant benefits and ethical complications (see *Nature* **469**, 289–291; 2011).

With the technology in place to sequence the fetal DNA carried in a pregnant woman’s bloodstream, geneticists predict the list of conditions that can be detected by non-invasive means will grow rapidly. Another company, Gene Security Network of Redwood City, California, says its forthcoming test will also check for other genetic abnormalities, and Sequenom is studying the feasibility of expanding its test.

“There’s every reason to think that in the future you’ll be able to extract an enormous amount of information from that sequencing data,” says Peter Benn, director of the Diagnostic Human Genetics Laboratories at the University of Connecticut Health Center in Farmington.

Sequenom’s test sequences 36-base-pair fragments of DNA to identify sections from chromosome 21. Normally, the chromosome contributes 1.35% of the total maternal and fetal DNA in the mother’s blood. An overabundance of this material indicates the genetic abnormality that marks Down’s syndrome.

Sequenom is marketing its test as an add-on to current screening methods, which estimate the chance that a woman is carrying a fetus with Down’s syndrome from ultrasound results and protein markers in the blood. Such non-genetic screening can detect 90–95% of Down’s syndrome cases, but falsely indicates that up to 5% of women are carrying a baby affected by the condition. Sequenom’s test could be taken after a positive screening result to help a woman decide whether to undergo amniocentesis, a test that extracts amniotic fluid with a needle and carries a small risk of miscarriage. A study published this month, and paid for by

Sequenom, found that the company’s test has a false positive rate of 0.2% (G. E. Palomaki et al. *Genet. Med.* <http://dx.doi.org/10.1097/GIM.0b013e3182368a0e>; 2011).

It could spare some women from having amniocentesis after a false-positive screening result. But Benn says that the test will also pose difficulties. For instance, because it would take 8–10 days to get the results of Sequenom’s test, if a woman did still opt for amniocentesis, and the result confirms that the baby has Down’s syndrome, there would be little time left to decide whether to terminate the pregnancy. And some women who test positive on MaterniT21 will probably choose to terminate pregnancies immediately rather than have amniocentesis.

“Inserting this new test in the way that Sequenom is proposing is very difficult, from the patient perspective, and difficult for physicians and counsellors to manage,” Benn says.

Ethicists also caution that using such easy screening methods ever earlier in pregnancy might worsen the gender imbalance seen in countries such as China and India. And if it becomes routine to check for many different kinds of genetic abnormalities, ethicists predict that more couples may face the quandary of whether to carry an ‘unhealthy’ fetus to term.

“The idea that couples have choices about whether to continue their pregnancies may become strained because parents may be seen as irresponsible for allowing ‘defective’ pregnancies to go to term,” says Mildred Cho, an ethicist at Stanford University in Palo Alto, California. Other ethicists worry that fears of eugenics will be raised if testing can be done for less-serious conditions.

Sequenom is solely focused on developing tests for conditions that are already part of prenatal screening programmes, says Mathias Ehrlich, the company’s senior director for research and development diagnostics. “We do not want to invent new applications. Our focus is on making existing clinical applications safer,” he says. “I don’t think that we are in a position to say that we should determine what hair colour the baby has.” ■

**“In the future you’ll be able to extract an enormous amount of information from that sequencing data.”**

## REGENERATIVE MEDICINE

# European ban on stem-cell patents has a silver lining

*Researchers can work without fear of action over patent infringement.*

BY EWEN CALLAWAY

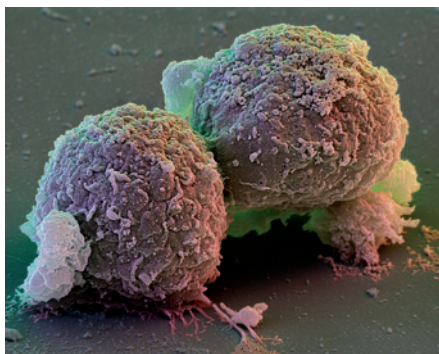
To hear European stem-cell researchers talk last week, you might have thought that their world was ending. After the European Court of Justice ruled on 18 October that procedures involving human embryonic stem (ES) cells cannot be patented, many responded with shock and dismay.

"This is the worst possible outcome and it's a disaster for Europe," Oliver Brüstle at the University of Bonn, Germany, told *Nature* shortly after learning that the court had felled his 1999 patent for a method of transforming human ES cells into neurons. Others said that without patent protection, few investors would pay to develop stem-cell therapies for conditions from neurodegenerative diseases to diabetes.

But in the days following the ruling, lawyers, funders and researchers have taken a more moderate view. There are other ways for companies and scientists who commercialize ES cells to protect their inventions in Europe, they say. And some believe that a lack of patents could speed up, rather than suffocate, innovation. "If anything the ruling is an opportunity," says physician scientist Chris Mason of University College London. "It's not the end of stem cells in Europe."

The decision by the European Court of Justice, which applies throughout the European Union and cannot be appealed, stems from a 2004 lawsuit brought by Greenpeace. The Amsterdam-based environmental group challenged Brüstle's patent on the grounds that it offended public sentiment and violated European law banning the industrial use of human embryos. A German court agreed, and by 2009 Brüstle's appeal had reached Europe's highest court (see *Nature* 462, 265; 2009). The language in these legal rulings — that commercial use of human embryos "would be contrary to ethics and public policy", for example — alarmed scientists, who spoke out against the court (A. Smith *et al.* *Nature* 472, 418; 2011).

The 13 judges of the court's Grand Chamber have now concluded that procedures involving human ES cells cannot be patented if they derive from the destruction of embryos. The ban applies retrospectively, and contrasts sharply with the position in the United States, where scientists face few restrictions on patents relating to ES-cell applications.



Embryonic stem cells: contrary to ethics and public policy?

"Time will tell how serious it's going to be," says Nick Bassil, an intellectual-property lawyer at Kilburn & Strode in London, who represents companies developing stem-cell therapies. He adds that it may take years for the European Patent Office, national patent offices and courts to interpret the ruling.

However, even a restrictive interpretation should allow companies to patent the technologies needed to turn human ES cells into treatments, rather than patenting procedures involving the cells themselves. "If the sum total of this market were some cell lines, I would be deeply, deeply worried," says Julian Hitchcock, a life-sciences lawyer at Field Fisher Waterhouse in London. Growth media, equipment and chemicals that help scientists to work with stem cells could all be patented in Europe without running afoul of the high court's ruling, he says. For instance, Peter Coffey at the Institute of Ophthalmology in London and his team are working with the drug giant Pfizer to develop a human-ES-cell-based treatment for macular degeneration, a progressive disease of the retina that causes blindness. Their patents cover the placement of their retinal cells in the eye, not the cells themselves.

Rob Buckle, a programme manager at Britain's Medical Research Council (MRC) in London, agrees that investors will find other ways to protect their intellectual property, and adds that the ruling will not affect the MRC's spending on ES-cell research.

The sheer complexity of therapies involving human ES cells should

also help to ward off copycats who might otherwise exploit the lack of patent protection to rush their own versions of a treatment to market. By keeping many of their manufacturing processes secret until they seek regulatory approval, companies can ensure that knock-offs are unlikely, says Mason. "If I give you my cell line, your chance of knowing what to do with it and copying what I do is zero," he says.

Many of the 20-year patents issued for ES-cell treatments will probably have expired by the time the treatments reach the clinic anyway, Mason adds. Indeed, the European Medicines Agency offers additional protection for inventions. The drug regulator keeps private for eight years any data that companies submit with their application for marketing approval, and blocks others from using this information for another two.

The ruling may even turn out to be a boon for European stem-cell science, says Mason, creating an anything-goes atmosphere that could attract scientists from abroad. Non-commercial research is generally exempted from patent infringement claims, but many patents cover the cells' use as research tools, creating uncertainty about which methods researchers are allowed to use, says Hitchcock.

A January statement from the Hinxton Group, an influential consortium of scientists and ethicists, had expressed concern that stem-cell biology was becoming so thick with broad patents that key areas of the field were being walled off from scientists and entrepreneurs. "With patents gone, it's much easier to do anything," says Mason. ■

*Additional reporting by Alison Abbott.*

## CORRECTION

The News story 'Angry words over East Asian seas' (*Nature* 478, 293–294; 2011) wrongly implied that *Climatic Change* took a defined stance on the position of China's border in the South China Sea. In fact, co-editor Michael Oppenheimer merely told the authors of a paper containing a contested map that the journal would make space for any amendments to the map that they may deem appropriate.

➔ **NATURE.COM**  
Insight on  
regenerative  
medicine:  
[go.nature.com/ax5uif](http://go.nature.com/ax5uif)



# Brain child

*Asking parents to donate a child's brain to research is emotionally fraught. Some researchers say that it is time to put aside the taboos.*

BY ALISON ABBOTT

David Amaral wanted to watch the young brain take shape. He thought that studying post-mortem brains under the microscope would help him to work out why children with autism often have abnormalities in the key structures that drive emotion and behaviour. But he soon found that existing brain banks couldn't give him what he needed. "It's just too hard to get high-quality tissue," he says. The banks may contain hundreds or even thousands of brains — but not from children, and not necessarily in the best condition.

Amaral, who is director of research at the MIND (Medical Investigation of Neurodevelopmental Disorders) Institute at the University of California, Davis, is not the only scientist eager for access to brains from children. The crucial stages of brain development span early fetal life through to the end of the teenage years; and destructive neurodevelopmental disorders such as autism and schizophrenia are thought to arise partly because of faulty connections laid down during this time. Many researchers want to apply new technologies, including increasingly sensitive molecular analyses and ever smarter microscopy, to developing brains to create a dynamic picture of what goes wrong.

When they succeed, the results can be breathtaking, says neuropathologist Joel Kleinman at the National Institute of Mental Health (NIMH) in Bethesda, Maryland. In work reported in this week's *Nature*<sup>1</sup>, he and his colleagues applied genomic technologies to 269 brains spanning the human lifetime and revealed an extraordinary wave of changes in gene expression that occur as the human brain develops. "It's like I witnessed the poetry of birth," he says.

But experiences such as Kleinman's are rare, owing to the challenges of collecting and storing children's brains. Parents must give permission shortly after their child has died, a time of inconsolable grief, and fetal

brains are available only after an abortion — an incendiary political issue as well as an emotionally painful one for the women involved. Biomedical organizations have been tiptoeing around the delicacies for a decade or more.

The solution, according to Amaral, is not complicated. Outreach programmes could be aimed at the coroners who conduct autopsies as well as at the families of children with brain disorders. They could explain the research value of donated brains and encourage families to sign up to a donor registry. A network of brain-collection centres around the United States could ensure that brains are preserved quickly. And centralized governance of the banks could direct tissue from each donated brain towards as much high-quality research as possible. "All it needs is for someone to take ownership of the issue," Amaral says.

That ownership may now be emerging from advocacy groups for neurodevelopmental disorders. "I know there has been a lot of talk and no action till now," says neuroscientist Robert Ring, vice-president of translational research at Autism Speaks, a research and advocacy organization based in New York. So Ring is pushing forward plans for a bank along the lines Amaral suggests. "Give us one year and we'll have developed a collaborative model with the scientific community," he says.

Only two major brain banks store brains from children or fetuses and distribute them to the research community at large. One is run by the National Institute of Child Health and Human Development (NICHD) and held at the University of Maryland School of Medicine in Baltimore; the other, called the Autism Tissue Program, is run by Autism Speaks and is hosted at the Harvard Brain Tissue Resource Center in Belmont, Massachusetts.

**➔ NATURE.COM**  
Read more about the  
adolescent brain at  
[go.nature.com/pzw4xd](http://go.nature.com/pzw4xd)

ILLUSTRATION BY GRACIA LAM

At most brain banks, including the NICHD's, personnel typically call the local coroner's office each morning. If a child is to be autopsied, they ask the office's permission to contact the family and request the brain for research. But the few coroner's offices involved can collect only a small amount of tissue. The Autism Tissue Program depends more on families that get in touch when they experience such a bereavement. Experts then go out to retrieve and prepare the brain. As the programme collects brains from across the United States, this often means a long journey. Ideally, though, the brain should be acquired quickly after death to minimize the breakdown of proteins and other molecules that researchers might wish to study. Other factors also influence tissue integrity, such as how soon after death a body is refrigerated and whether the person died slowly and painfully, as scientists have shown that this alters gene expression in the brain, making it less useful for research.

Collecting fetal brains is also hard. Brains from spontaneous abortions can't be used for research because the fetus has generally been dead for many hours before it is expelled. In fact, brains can be collected from abortions only when labour has been induced medically, because surgical procedures tend to damage the tissue.

Neither the NICHD bank nor the Autism Tissue Program bank — which together hold nearly 1,300 brains from people aged 19 and under — can meet the demand from researchers. Neuroscientist H. Ronald Zielke, director of the NICHD bank, says that he turns down 20% of requests for tissue because of a lack of material. In particular, this and other brain banks are running critically short — or have run out — of the brain areas that are the most interesting for research into developmental disorders, says Zielke. That includes the amygdala, which processes emotion, and the prefrontal cortex, which processes other cognitive and social behaviours. A brain bank, like any tissue repository, is also very expensive to run — the annual direct costs for the NICHD bank come to US\$900,000.

To get around the shortage, some researchers have built up collections for their own use. Kleinman's research on gene expression drew on a collection that he heads at the NIMH. A similar study in this week's *Nature*<sup>2</sup>, led by Nenad Šestan from the Yale University School of Medicine in New Haven, Connecticut, and with Kleinman as a co-author, drew in part on a collection that Šestan has generated at Yale. Their study showed the dramatic changes in gene expression that occur before and shortly after birth (see 'Brain waves'). Neonatologist David Rowitch at the University of California, San Francisco, began a collection of brains at his hospital, which led to a paper published in last week's *Nature*<sup>3</sup> showing that the migration of 'progenitor' cells between two brain structures seen in infants slows down after the age of 18 months and has almost disappeared by adulthood. He began collecting brains in 2008 with the support of the Howard Hughes Medical Institute in Chevy Chase, Maryland, and now has more than 100, most of which are from very young babies.

These studies show how valuable such collections can be, but both Rowitch and Šestan describe the process of creating and running their own banks as "a big headache" because of the bureaucracy associated with handling human material. Šestan says that he would feel "much more comfortable" if the National Institutes of Health (NIH) were to run his collection. "It's a huge effort for a small group and the NIH could do something on a larger scale," he says.

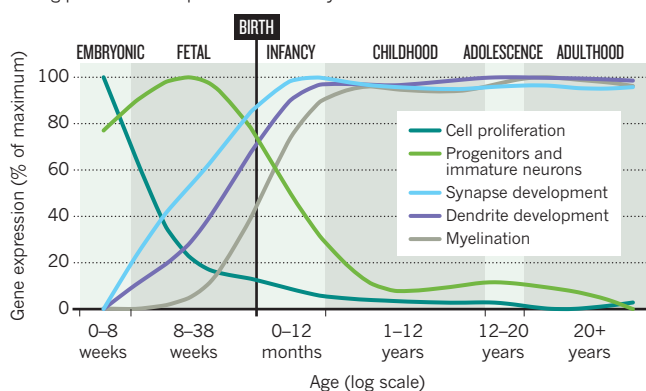
## PUTTING BRAINS TOGETHER

In fact, neuroscientists have been proposing for years that the NIH take a leading role in establishing a network of collection centres and standardizing methods for brain collection and preservation.

In July last year, Autism Speaks and the other major US foundation that funds autism work, the Simons Foundation in New York, made a formal proposal to the NIMH for a public-private partnership to collect brains from children with and without autism. The idea is that the advocacy groups would engage in intensive outreach efforts to potential

## BRAIN WAVES

The activity of genes linked to neural development changes most dramatically during prenatal development and infancy.



donors, particularly families who have a child with autism, and the NIH would fund and manage the bank.

The NIH, though, has been slow to commit. Ring, who moved from the drug giant Pfizer to Autism Speaks in June this year and has the candor of someone used to industry deadlines, sees "a unique opportunity for the foundations to take on a leadership role". His organization and the Simons Foundation are now in discussions with scientists to get agreement on scientific standards for the bank. He says that multiple collection centres will help to overcome geographical logistics, shortening the time from death to collection, for example.

Thomas Insel, director of the NIMH, says that the NIH already supports 11 brain banks related to different neurological disorders, and would like to adopt "a rational overall strategy rather than simply adding another boutique brain bank to the list". He says that the NIH has now agreed in principle, at least, to create a 'neurobiobank' that would include both adult and children's brains. Although no firm plans have been released, the bank would probably have multiple collection points (the agency's existing tissue banks would become 'nodes'), but centralized oversight and tissue distribution. That is essentially what the advocacy groups want.

However the banks are organized, the agonizing task of approaching bereaved families will remain. Yet autism researcher Cynthia Schumann, who earlier this year became director of an effort by the MIND Institute to start a bank of its own, says that her first encounters with families who choose to donate were eye-opening. "I have been blown away by how parents have thanked us — for helping them to handle grief with the opportunity to give something back to help autism research," she says. Schumann, like counsellors at Autism Speaks, has also spent time educating affected families about autism research. "Parents often agree to sign up to a registry, and to encourage other families to sign up too," she says. So the reluctance to ask parents about acquiring their children's brains, she thinks, may be ill-founded.

That seems to be reflected in the experience of Valerie Hund, who donated the brain of her 16-year-old son, Grayson, to the MIND Institute after he died in January. Grayson had autism and epilepsy, and had died during a seizure. Hund says that a neighbour was a board member of the MIND Institute, and that her elder daughter had thought to call him shortly after Grayson died. The donation, says Hund, "helped me to cope through the process. I'm happy that Grayson is a pioneer in this."

Hund says she thinks that the institute's programme for raising awareness on brain and tissue banking is important. "It would have been easier for us if we had thought about donation in advance — but that is the last thing on your mind." ■ SEE EDITORIAL P.427

Alison Abbott is *Nature's* senior European correspondent.

1. Colantuoni, C. *et al. Nature* **478**, 519–523 (2011).
2. Kang, H. J. *et al. Nature* **478**, 483–489 (2011).
3. Sanai, N. *et al. Nature* **478**, 382–386 (2011).



# The Black Death DECODED

THE GENOME OF A 660-year-old bacterium is revealing secrets from one of EUROPE'S DARKEST CHAPTERS.

By Ewen Callaway



A word of a brutal pestilence raging across Europe reached London, its residents started digging. In 1348, Ralph Stratford, Bishop of London, dedicated acres of land that had been purchased to bury the legions of Black Death victims who would overwhelm existing churchyard cemeteries. Within two years, one-third to one-half of the city's 40,000–100,000 residents succumbed, and many thousands were buried in two newly dug cemeteries at East and West Smithfield. At the height of the scourge, 200 bodies were interred each day.

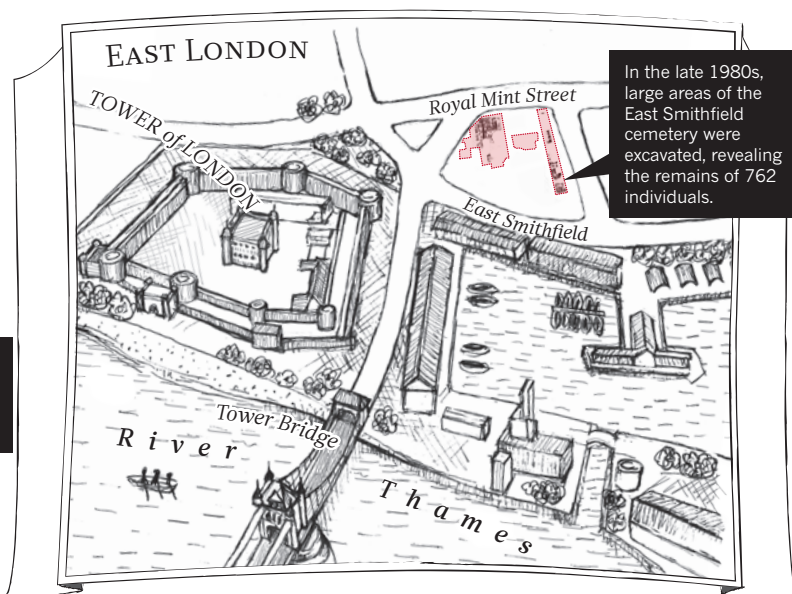
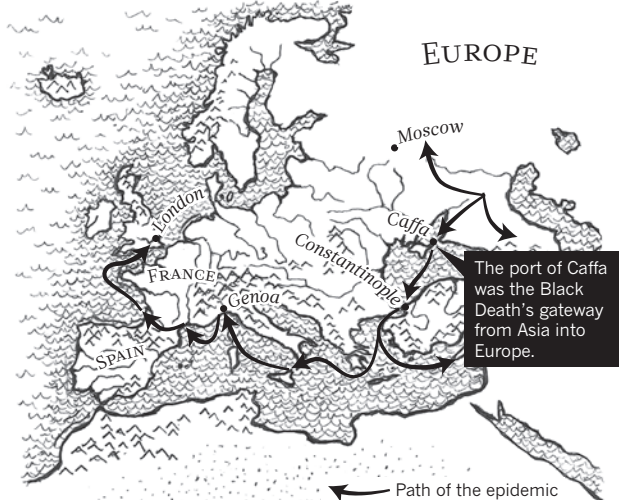
East Smithfield, originally called the Churchyard of the Holy Trinity, is one of a handful of burial sites known to have been used only during the Black Death. In the 1980s, excavation of this 'plague pit' turned up nearly a third of the 2,400 bodies estimated to be buried there, some piled five deep. Despite the urgency of the time, the bodies were placed purposefully, oriented east to west, some with charcoal, possibly to absorb the fluids released during putrefaction, and many with coins and trinkets of their former lives. Such foresight not only helped keep corpses from piling up in the streets, but also, it seems, afforded some Black Death victims a dignified Christian burial. Six-and-a-half centuries later, it would also give scientists the opportunity to dissect the disease that laid waste to Europe (see 'Death on the march').

This month, geneticists reported that they have reconstructed the genome of *Yersinia pestis*, the bacterium that causes bubonic plague, recovered from remains at East Smithfield<sup>1</sup>.

SOURCE: REF. 1

## DEATH ON THE MARCH

In the 1340s, a pestilence originating in Western Asia spread rapidly across Europe. Before it overtook London in 1348, land was set aside in East Smithfield to bury the dead.



The sequence — the first from an ancient bacterial pathogen — may help to explain how a disease could wreak so much havoc. It also marks a renaissance in genetic studies of ancient diseases, a field that has suffered a controversial history but that is now being revitalized. “There will be a race now for all the ancient pathogens,” says Hendrik Poinar, a palaeogeneticist at McMaster University in Hamilton, Canada, who co-led the sequencing efforts.

### PLAGUED WITH DISBELIEF

When Alexandre Yersin linked *Y. pestis* to bubonic plague in 1894, many scientists surmised that the pathogen was behind not only the Black Death, but also a spate of earlier mass die-offs. The sixth-century Justinian plague devastated Constantinople and killed millions in Europe and the Near East. Plagues reared their heads periodically for the next two centuries. Black Death itself reappeared several times, even into the nineteenth century.

Clues tying *Y. pestis* to these outbreaks came largely from historical accounts of their symptoms, such as Giovanni Boccaccio’s description of the Black Death in *The Decameron*, written around 1350: “It first betrayed itself by the emergence of certain tumours in the groin or the armpits, some of which grew as large as a common apple, others as an egg.”

But some modern historians and scientists came to doubt that *Y. pestis* caused these ancient outbreaks. Bubonic plague epidemics known to have been caused by *Y. pestis* in the past century seemed too mild to have been caused by the same culprit as the Black Death: they killed fewer people and spread more slowly. Some ‘plague revisionists’ have argued that fleas, which spread *Y. pestis* to humans, would have struggled to survive the cold temperatures reported during the Black Death. And there was the speed with which it killed — Boccaccio reported that death often occurred within three

days of the first symptoms appearing. Anthrax or a haemorrhagic-fever-causing virus similar to Ebola would be more likely than plague to cause such a rapid demise, say critics.

DNA evidence would seem to offer a definitive answer. In 2000, a team led by Didier Raoult, a microbiologist at the University of the Mediterranean in Marseilles, France, said it had proved the link between the bacterium and the disease. The researchers reported<sup>2</sup> that they had successfully recovered *Y. pestis* DNA from the teeth of a child and two adults dug up from a fourteenth-century mass burial site in Montpellier. The team identified the bacterium using a sensitive technique called the polymerase chain reaction (PCR) to amplify a portion of a gene from *Y. pestis* called *pla*. “We believe that we can end the controversy,” the team wrote<sup>2</sup>. “Medieval Black Death was plague.”

But several critics raised concerns about contamination. The PCR might instead have amplified DNA from modern *Y. pestis* used previously in the lab, or possibly the sequences from a closely related soil-dwelling bacterium. “I could never, ever replicate it,” says Thomas Gilbert, an evolutionary geneticist at the University of Copenhagen in Denmark. In 2004, Gilbert and his colleagues reported no trace of *Y. pestis* DNA in 108 teeth from 61 individuals found in plague pits in France, Denmark and England (including East Smithfield)<sup>3</sup>.

Raoult says that there was no contamination and that Gilbert’s methods did not accurately replicate his<sup>4</sup>. Still, those who were already sceptical of the suggestion that *Y. pestis* caused the Black Death latched on to Gilbert’s study.

Other studies of microbial DNA extracted from ancient human remains — including those affected by tuberculosis, syphilis and malaria — were also being scrutinized. In several cases, researchers could not replicate results, or they found methodological shortcomings. Critics said that DNA from these

samples was too degraded by heat, moisture and time to detect, and the field soon divided into believers and sceptics.

“There was a complete schism,” says Ian Barnes, a palaeogeneticist at Royal Holloway University of London, who says he spent two-and-a-half years trying — unsuccessfully — to find DNA evidence of syphilis or tuberculosis in bones dating from the nineteenth and early twentieth centuries<sup>5</sup>. “People largely ignored each other,” he says.

### DIGGING UP ANSWERS

Although Poinar was dubious of claims about ancient microbial DNA, he was intrigued by the bones from East Smithfield. Nearly all of the remains are from Black Death victims, many of whom were cut down during the prime of their lives.

In a bright ground-floor laboratory of the Museum of London, a short walk from East Smithfield, osteoarchaeologist Jelena Bekvalac examines the nearly complete skeleton of one of the plague pit’s former residents. Wearing a black silk scarf dotted with white skull-prints, Bekvalac handles a pelvic bone and determines that it belonged to a man who died in his late teens or early twenties. Apart from some plaque on his teeth and a gash in his skull that shows some signs of healing, the man’s skeleton offers no outward evidence of Black Death.

His remains, and those from hundreds of others, represent a snapshot of life and death in London during the epidemic. Since the site’s excavation, researchers have descended on the bones in search of information.

In the late 1990s, Poinar met Sharon DeWitte, then a graduate student at Pennsylvania State University in State College, who was working on a demographic analysis of the remains suggesting that Black Death preferentially killed those who were already frail. The two considered drilling into teeth

and bones to find *Y. pestis* DNA, but Poinar wasn't satisfied with the available detection tools, which were still based on PCR. "We sort of sat on the samples for a few years waiting for all the stars to align," says DeWitte, now at the State University of New York at Albany.

That alignment came from next-generation DNA sequencers, machines that read short snippets of DNA. The technology was perfect for sequencing DNA that has been damaged by spending hundreds of years underground.

The sequencers allowed Svante Pääbo, a palaeogeneticist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and his team to sequence a draft of the Neanderthal genome<sup>6</sup>. But finding and sequencing ancient pathogens in a human skeleton is much harder — like finding "needles in the football field", Poinar says — because their genomes are 1,000 times shorter than that of the Neanderthal and closely resemble those of soil microbes that have infiltrated the bones.

Another technology helped narrow the search. Pääbo and his team developed a technique, called targeted capture, in which they used lab-synthesized 'bait' DNA to snag ancient DNA strands from a bone sample<sup>7</sup>, leaving soil-microbe and other sequences behind. "It's pretty much like fishing in a pond," says Johannes Krause, a palaeogeneticist at the University of Tübingen in Germany, who worked with Pääbo on the Neanderthal genome and co-led the Black Death project with Poinar.

In a proof-of-principle experiment published in August of this year, Krause and Poinar's team used sequences from a contemporary plague strain to fish out *Y. pestis* DNA from the teeth of victims buried at East Smithfield. From this, they sequenced a short loop of DNA, called the pPCP1 plasmid, that is partially responsible for bubonic plague's ability to infect humans.

Their results<sup>8</sup>, along with a paper published last year<sup>9</sup> that found *Y. pestis* sequences in different Black Death bone samples, have convinced most scientists that bubonic plague was involved in the Black Death.

In their most recent paper<sup>1</sup>, Poinar and Krause completed the ancient genome and showed that it sits at the root of an evolutionary tree that comprises 17 contemporary strains of *Y. pestis*. This indicates that the Black Death strain spawned many of the forms of *Y. pestis* that infect humans today.

This strain, Krause adds, probably emerged not long before the Black Death started its rampage across western Asia and Europe in the fourteenth century. "That, for me, was the biggest surprise," he says. It suggests, the authors argue, that earlier plagues were caused by either a now-extinct strain of *Y. pestis* or by an entirely different pathogen.

Mark Achtman, a plague-evolution expert at University College Cork in Ireland, calls this interpretation "absolute

nonsense". Krause and Poinar's team did not consider a number of modern plague strains found in central and east Asia, which are thought to have earlier origins than the East Smithfield strain, Achtman says. Genome sequences for these strains were not available to his team, says Krause, but he is eager to see how they are related.

### MYSTERIOUS SCOURGE

Just as puzzling, however, is that *Y. pestis* seems to have changed very little over the past 660 years. The genome of the Black Death strain differs from that of the modern *Y. pestis* 'reference' strain by about 100 nucleotides, but each of these genetic differences can be found in at least one contemporary strain. "We



Historical descriptions of the Black Death have helped link *Yersinia pestis* with the disease.

can't find anything that makes the Black Death special," Krause says.

The team is now looking for other genetic changes that could account for the Black Death's ferocity, such as rearrangements in the genome, which are difficult to determine from the short fragments of DNA available. To better understand how the plague worked, researchers could try to resurrect the Black Death pathogen by modifying the genomes of contemporary *Y. pestis* strains. Although this might sound alarming, research on *Y. pestis* is already carefully controlled, and even an accidental infection with such a strain could be easily treated with modern antibiotics.

Moreover, Poinar says, the Black Death was not just about the bacterium. Environmental and epidemiological factors must have aided in its vicious tear through Europe. Sick soldiers returning to Europe from Caffa, the Black Sea port that was the plague's gateway from Asia, unleashed the disease on a population that would have been weakened by malnourishment and years of cold, wet weather, he says.

Achtman says that it is possible that Black Death was not spread by rat-dwelling fleas, as *Y. pestis* is today, but by other animals, which could have enhanced transmission. Or another

circulating pathogen could have contributed, as in the 'Spanish flu' pandemic that killed up to 100 million people worldwide in 1918-19, often with the help of bacterial pneumonia.

Whatever questions remain about the Black Death, scientists are now keen to apply the latest sequencing methods to other ancient epidemics. "I've completely gone from thinking, 'ancient pathogens are a load of crap,' to 'hold on, maybe some of this stuff works,'" says Gilbert, whose team has started to sequence DNA from pathogens that plagued ancient crops. Researchers could identify ancient microbes and chart their spread and their evolutionary relationships with contemporary strains. For example, Europeans who travelled to the New World may have introduced new forms of tuberculosis to North America and brought syphilis back to Europe.

Ancient pathogens may help scientists understand current and future outbreaks, says Terry Brown, a biomolecular archaeologist at the University of Manchester, UK. He and Charlotte Roberts, of Durham University, UK, are charting the evolution of tuberculosis strains in Britain and Europe. "By looking over the past 1,000 years of disease in British cities, we can understand problems occurring in the Third World, where more and more people are crowding into cities," he says. Similarly, the sequencing and resurrection of the influenza strain responsible for the 1918 pandemic<sup>10</sup> has helped researchers to interpret the sequences of contemporary flu strains.

For all its ferocity, the Black Death left few visible marks on London. Today, the plague pit at East Smithfield is in the heart of London's financial district, buried under modern office suites and the old Royal Mint building. The only visible remnants are the crumbled ruins of St Mary Graces, a Cistercian abbey built near the site in 1350.

London may have seen its last significant bubonic plague outbreak, but catastrophic epidemics are a rule of human history, not an exception. Centuries from now, what traces will the next great scourge leave? Future archaeologists chronicling its history may find memorials, graves and probably even the bodies of victims. But another story will also lurk in its DNA, just waiting to be read. ■

Ewen Callaway writes for Nature from London.

1. Bos, K. I. et al. *Nature* **478**, 506–510 (2011).
2. Raoult, D. et al. *Proc. Natl Acad. Sci. USA* **97**, 12800–12803 (2000).
3. Gilbert, M. T. P. et al. *Microbiology* **150**, 341–354 (2004).
4. Drancourt, M. & Raoult, D. *Microbiology* **150**, 263–264 (2004).
5. Barnes, I. & Thomas, M. G. *Proc. R. Soc. B* **273**, 645–653 (2006).
6. Green, R. E. et al. *Science* **328**, 710–722 (2010).
7. Briggs, A. W. et al. *Science* **325**, 318–321 (2009).
8. Schuenemann, V. J. et al. *Proc. Natl Acad. Sci. USA* **108**, E746–E752 (2011).
9. Haensch, S. et al. *PLoS Pathog.* **6**, e1001134 (2010).
10. Taubenberger, J. K. et al. *Nature* **437**, 889–893 (2005).

# COMMENT

**CLIMATE CHANGE** Future droughts will threaten food security **p.450**

**SOCIOLOGY** Scientists mark their tribes with tattoos **p.454**



**MATHEMATICS** Persi Diaconis on magic tricks and card shuffles **p.457**

**OBITUARY** Ralph Steinman, the discoverer of dendritic immune cells **p.460**



Families in Bangladesh seek safer areas after severe floods in 2007.

## Migration as adaptation

Mobility can bring opportunities for coping with environmental change, say **Richard Black, Stephen R. G. Bennett, Sandy M. Thomas and John R. Beddington.**

The effects of global environmental change, including coastal flooding, reduced rainfall in drylands and water scarcity, will almost certainly alter patterns of human migration. Conventional narratives usually cast these displacements in a negative light, with many millions of people forced to move, and tension and conflict the result. Our study suggests that the picture is not so one-sided.

The study, the UK government's Foresight report on migration and global environmental change, examines the likely movement of people within and between countries over the next 50 years<sup>1</sup>. It contends that, although environmental change will alter an already complex pattern of human mobility, migration will offer opportunities as well as challenges. The greatest risks will be borne by those who are unable or unwilling to relocate, and may be exacerbated by maladaptive policies designed to prevent migration. It is time for a fresh discourse — and fresh research — on migration in relation to global environmental change.

International action and research are needed to identify the positive and negative outcomes of migration influenced by environmental change. Whether movement occurs within or between countries, there is a need to prepare for it and in some cases enable it. It is important to deepen understanding of how migration will affect other types of social change, such as the evolution of cities, the formation of 'poverty traps' and the coexistence of cultures. Current policy frameworks should take account of these factors to avoid having to deal later with impoverishment and displacement under high-risk conditions.

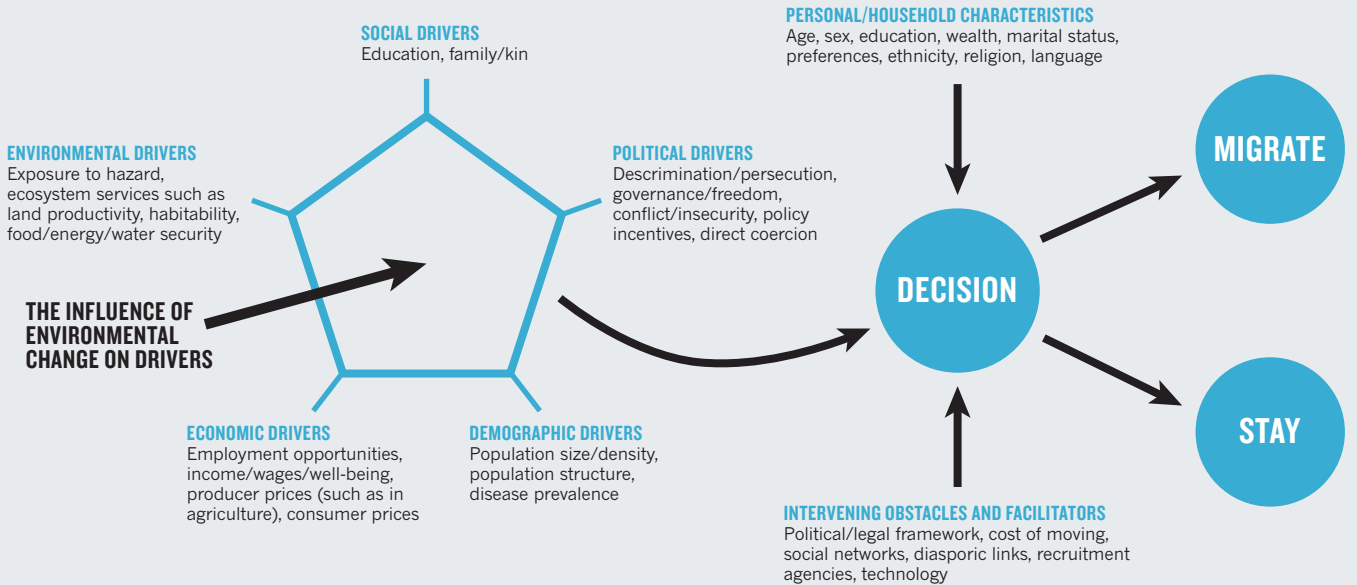
### THE REALITY OF MIGRATION

Many people across the world are already migrating, motivated by strong socio-economic factors. The United Nations estimates that there are about 210 million international migrants, but as many as 740 million internal (intranational) migrants<sup>2</sup>. People migrate for complex reasons: to improve incomes; to join family members; to escape persecution; and to remove themselves from environmental or other threats, often temporarily. Such

**NATURE.COM**  
Migration: an engine for social change:  
[go.nature.com/lebnix](http://go.nature.com/lebnix)

## THE DRIVERS OF MIGRATION

Many factors influence whether a person or family will migrate. Their effects are closely intertwined, so it makes little sense to consider any of them in isolation.



► drivers will change in their proportions over coming decades, but how they influence people's decisions about where they live will not (see 'The drivers of migration').

Environmental factors will increasingly influence migration. Current greenhouse-gas emissions are already committing the planet to likely climate changes in the next 20 years. In Bangladesh, for example, moving to cities has become a common coping strategy in the face of flooding. In a 2008 study, 22% of households affected by tidal-surge floods, and 16% affected by riverbank erosion, moved to urban areas<sup>3</sup>. Diminished soil quality in Kenya has led people to travel to diversify their income. In 2004–05, for example, temporary labour migration in households that were farming land with high-quality soils was 67% lower than in those using poor soils<sup>4</sup>.

Not everyone is able to migrate. There may be confounding socio-political factors, such

as in Somalia where armed conflict restricts movement<sup>5</sup>, or in New Orleans, Louisiana, where the evacuation plan for Hurricane Katrina assumed that everyone had access to a car<sup>6</sup>. Migration is often expensive, and those most vulnerable to environmental change are usually poor. For example, in Uganda, a relatively settled country with high 'entry costs' for housing, schools and marriage, those who are wealthier are more able to relocate. In Mali, emigration decreased during the severe droughts of 1983–85 alongside a rise in rural poverty<sup>7</sup>.

### INTERLINKED FACTORS

Environmental change can increase the incentive to move, but it can also limit the capacity to do so. It should be seen as affecting the many linked drivers of migration.

People are as likely to migrate into places of environmental vulnerability as away from

them — a point that has been insufficiently acknowledged. In rapidly growing megacities, such as Dhaka and Lagos, that are located in delta and coastal floodplain regions in Africa or Asia, hundreds of millions more people may be at risk of flooding by 2060 (see 'Urban coastal flood risk'). Migrants stretch the capacity of existing infrastructure, especially in low-income countries, and new arrivals are frequently vulnerable. In Dakar, Senegal, for example, 40% of those who moved there between 1998 and 2008 live in areas of high flood risk<sup>8</sup>.

Political instability, poor governance, conflict and social pressures compound these problems. For example, Zimbabwe's political and economic crisis, amplified in rural areas by drought, has contributed to the migration of between 1.5 million and 2 million Zimbabweans to South Africa since 2000. In 2008, attacks against these migrants resulted in 65 deaths and the further displacement of 150,000 people<sup>9</sup>.

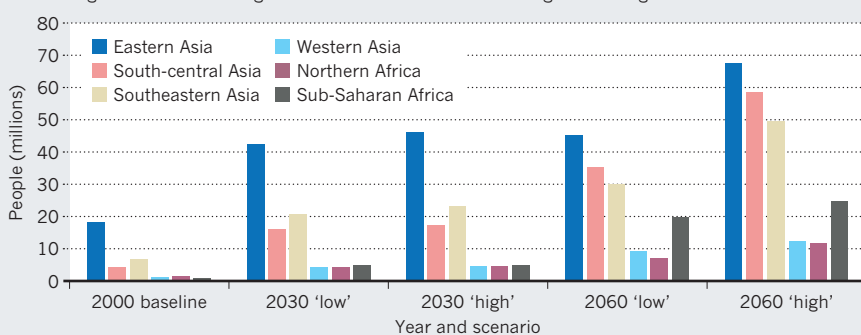
In many cases of mass migration, especially when coupled with environmental hazard, humanitarian assistance might be needed. And such upheavals may have political ramifications. If sea level rise were to engulf a small island state, for example, it would raise issues of sovereignty, and questions of who is responsible for displaced populations.

Migration may be the most effective way to allow people to diversify income and build resilience where environmental change threatens livelihoods. It is therefore necessary to make channels for voluntary migration available.

Within countries, this implies removing arbitrary restrictions on movement, and

## URBAN COASTAL FLOOD RISK

The number of people living in cities that are at risk of coastal flooding is set to increase dramatically over the coming decades in both 'high' and 'low' scenarios of economic growth and governance\*.



\*Low scenario = high economic growth and inclusive governance; high scenario = low global economic growth and fragmented governance



Poverty trap: residents of this slum in Luanda would find it hard to avoid environmental changes.

providing basic infrastructure to enable relocation and settlement in urban areas, ideally sustainably. Internationally, this might include the extension of regional economic communities to cover the free movement of people as well as money and goods. Those at risk of being trapped — the poorest and least mobile — require additional measures, such as functional early-warning systems and tested emergency evacuation plans, to minimize their vulnerability to extreme events (see ‘Trapped populations’).

### A ROAD MAP FOR ACTION

For international policy-makers, climate mitigation and the reduction of negative environmental change should continue to be priorities. But mechanisms for funding adaptation to climate change also need to account for migration as a way of building resilience. It is therefore important that long-term initiatives, including for example those instigated under the United Nations Framework Convention on Climate Change, recognize the links between global environmental change and migration.

These initiatives should consider the realities of migration, including benefits as well as challenges. And they should focus on the resilience of populations that are moving to, or are trapped in, urban areas that are vulnerable to environmental change, particularly in low-income countries.

Other actions to boost resilience — including sustainable urbanization, climate-smart development, conflict resolution and emergency preparedness — need to take account of an increased propensity for people to migrate. Planners will need to provide

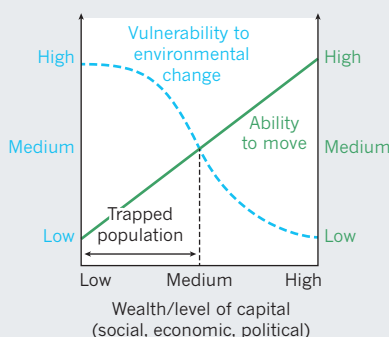
flood-control, water-management, forecasting and warning capacities to growing urban populations. Furthermore, migrants may be socially excluded and so will need special attention.

To increase the potential benefits of international migration, policies can link its adaptive advantages for some migrant communities to demographic deficits and labour shortages in potential host nations. Circular migration schemes are one option, to allow those in environmentally vulnerable areas to work seasonally or on a temporary basis in countries where their skills are in demand. A strategic international approach to migration also needs to pay attention to regional and global demand for skilled workers in particular sectors.

Whole populations need not abandon

### TRAPPED POPULATIONS

Impoverished people face a double set of risks. They are unable to move away from environmental threats, and their lack of capital makes them especially vulnerable to environmental changes.



their former homes. Migration of some individuals can help a community to remain viable in the long run, for example if money and goods are sent back to help build their resilience. In Africa, where the majority of international migrants stay within their subregion, such remittances to home communities quadrupled to nearly US\$40 billion between 1990 and 2010, surpassing official development assistance since 2007 (ref. 10).

Researchers in the fields of development, climate and environmental science, and climate adaptation need to pay more attention to migration. A better understanding is required of the extent to which migration influences vulnerability and resilience in the face of environmental change. So, too, is clarity on the adequacy of policy responses to address the impact of global environmental change on migrant and non-migrant communities. Such knowledge must be based on empirical research, and underpinned by longitudinal data on migration flows.

It is vital for the research community to provide insights into what outcomes can be expected, and the Foresight project provides a framework with which to start this endeavour. ■

**Richard Black** is head of the School of Global Studies, University of Sussex, Falmer, Brighton BN1 9SJ, UK. **Stephen R. G. Bennett** is at the Foresight Programme, Government Office for Science, London SW1H 0ET, UK; **Sandy M. Thomas** is head of Foresight, Government Office for Science. **John R. Beddington** is chief scientific adviser to the UK government, and head of the Government Office for Science. e-mail: foresightweb@bis.gsi.gov.uk

1. Foresight: Migration and Global Environmental Change Final Project Report (Government Office for Science, 2011); available at <http://go.nature.com/somswg>
2. UNDP Human Development Report 2009: Overcoming Barriers: Human Mobility and Development (UNDP, 2009).
3. Penning-Rowsell, E., Sultana, S. & Thompson, P. Population Movement in Response to Climate Related Hazards in Bangladesh: The ‘Last Resort’ (Government Office for Science, 2011); available at <http://go.nature.com/somswg>
4. Gray, C. *Glob. Environ. Change* **21**, 421–430 (2011).
5. Kolmannskog, V. *New Issues in Refugee Research* Res. Pap. 180 (UNHCR, 2009); available at <http://go.nature.com/pugh4r>
6. Cutter, S. *The Katrina Exodus: Internal Displacements and Unequal Outcomes* (Government Office for Science, 2011); available at <http://go.nature.com/somswg>
7. Findley, S. E. *Int. Migr. Rev.* **28**, 539–553 (1994).
8. World Bank World Development Report 2010: Development and Climate Change (World Bank, 2010).
9. McGreggor, J., Marazzi, L. & Mpofu, B. *Conflict, Migration and the Environment: the Case of Zimbabwe* (Government Office for Science, 2011); available at <http://go.nature.com/somswg>
10. Ratha, D. et al. *Leveraging Migration for Africa: Remittances, Skills, and Investments* (World Bank, 2011).



Hundreds of thousands fled the 1930s US Dust Bowl; more drought-spurred migrations are expected.

# The next dust bowl

Drought is the most pressing problem caused by climate change. It receives too little attention, says  
**Joseph Romm.**

Which impact of anthropogenic global warming will harm the most people in the coming decades? I believe that the answer is extended or permanent drought over large parts of currently habitable or arable land — a drastic change in climate that will threaten food security and may be irreversible over centuries.

A basic prediction of climate science is that many parts of the world will experience longer and deeper droughts, thanks to the synergistic effects of drying, warming and the melting of snow and ice.

Precipitation patterns are expected to shift, expanding the dry subtropics. What precipitation there is will probably come in extreme deluges, resulting in runoff rather than drought alleviation. Warming causes greater evaporation and, once the ground is dry, the Sun's energy goes into baking the soil, leading to a further increase in air temperature. That is why, for instance, so many temperature records were set for the United States in the 1930s Dust Bowl; and why, in

2011, drought-stricken Texas saw the hottest summer ever recorded for a US state. Finally, many regions are expected to see earlier snowmelt, so less water will be stored on mountain tops for the summer dry season. Added to natural climatic variation, such as the El Niño–La Niña cycle, these factors will intensify seasonal or decade-long droughts. Although the models don't all agree on the specifics, the overall drying trends are clear.

I used to call the confluence of these processes 'desertification' on my blog, Climate-Progress.org, until some readers pointed out that many deserts are high in biodiversity, which isn't where we're heading. 'Dust-bowlification' is perhaps a more accurate and vivid term, particularly for Americans — many of whom still believe that climate change will only affect far-away places in far-distant times.

Prolonged drought will strike around the globe, but it is surprising to many that it

would hit the US heartland so strongly and so soon.

The coming droughts ought to be a major driver — if not the major driver — of climate policies. Yet few policy-makers and journalists seem to be aware of dust-bowlification and its potentially devastating impact on food security. That's partly understandable, because much of the key research cited in this article post-dates the 2007 *Fourth Assessment Report* of the Intergovernmental Panel on Climate Change (IPCC). Raising public awareness of, and scientific focus on, the likelihood of severe effects of drought is the first step in prompting action.

## AMERICAN NIGHTMARE

I first heard of the risks in a 2005 talk by climatologist Jonathan Overpeck of the University of Arizona in Tucson. He pointed to emerging evidence that temperature and annual precipitation were heading in opposite directions over many regions and raised the question of whether we are at the "dawn of the super-interglacial drought".

The idea wasn't new. As far back as 1990, scientists at NASA's Goddard Institute for Space Studies in New York projected that severe to extreme drought in the United States, then occurring every 20 years or so, could become an every-other-year phenomenon by mid-century<sup>1</sup>.

Events are starting to bear out these worrying predictions. Snowpack reduction, early snowmelt and a decrease in dry-season river flow in the American West, forecast more than two decades ago, have now been measured<sup>2</sup>. In much of the northern Rockies, the peak of the annual stream runoff is up to three or four weeks earlier than it was half a century ago<sup>3</sup>. Heat and drought — coupled with the greater impact of destructive species, such as bark beetles, aided by warming — have increased forest die-off and the risk of wildfire.

The palaeoclimate record dating back to the medieval period reveals droughts lasting many decades. But the extreme droughts that the United States faces this century will be far hotter than the worst of those: recent decades have been warmer than the driest decade of the worst drought in the past 1,200 years<sup>4</sup>.

And much warmer conditions are projected. According to a 2009 report of the US Global Change Research Program<sup>5</sup>, warming over mid-latitude land masses, such as the continental United States, is predicted to be higher than the forecast average global warming: much of the inland United States faces a rise of between 5°C and 6°C on the current emissions path (that is, 'business as usual') by the century's end, with a substantial fraction of that warming occurring by mid-century.

A 2007 analysis of 19 climate projections estimated that levels of aridity comparable to those in the Dust Bowl could stretch from

**NATURE.COM**  
Mega-drought threat  
to US Southwest:  
[go.nature.com/my4gcv](http://go.nature.com/my4gcv)



Then and now: sun-baked dry soils kick up clouds of dust in the 1930s (left) and in the modern United States (right).

Kansas to California by mid-century<sup>6</sup>. To make matters worse, the regions at risk of reduced water supply, such as Nevada, have seen a massive population boom in the past decade. Overuse of water in these areas has long been rife, depleting groundwater stores.

Of course, the United States is not alone in facing such problems. Since 1950, the global percentage of dry areas has increased by about 1.74% of global land area per decade<sup>7</sup>. Recent studies have projected 'extreme drought' conditions by mid-century over some of the most populated areas on Earth — southern Europe, south-east Asia, Brazil, the US Southwest, and large parts of Australia and Africa<sup>8</sup>. These dust-bowl conditions are projected to worsen for many decades and be "largely irreversible for 1,000 years after emissions stopped"<sup>9</sup>.

The concept of drought has not been ignored by the IPCC and other scientific groups; there is even a United Nations Convention to Combat Desertification. But the cumulative risks don't seem to have been fully recognized by the public and by policy-makers. And key questions remain to be answered, ideally in a dedicated report by an organization such as the US National Academy of Sciences or the IPCC.

## UNANSWERED QUESTIONS

Most pressing, what will happen to global food security if dust-bowl conditions become the norm for both food-importing and food-exporting countries? Extreme, widespread droughts will be happening at the same time as sea level rise and salt-water intrusion threaten some of the richest agricultural deltas in the world, such as those of the Nile and the Ganges. Meanwhile, ocean acidification, warming and overfishing may severely deplete the food available from the sea.

What are the implications of dust-bowlification for energy generation? After

agriculture, energy generation is responsible for the majority of freshwater withdrawals, and two key strategies for generating additional potable water — wastewater purification and desalinization — are both energy intensive. Future energy systems will need to be low on greenhouse-gas emissions and on water use. In particular, thermal power plants — including nuclear — may need to switch from evaporative or 'wet cooling'

**"Aridity comparable to the 1930s Dust Bowl could stretch from Kansas to California by mid-century."**

systems to dry cooling techniques, which, unfortunately, tend to be less efficient. From an ecological perspective, what will be the effects of dust-bowlification on the global carbon cycle? In the past six years, the Amazon has seen two droughts of the sort expected once in 100 years, each of which may have released as much carbon dioxide from vegetation die-off as the United States emits from fossil-fuel combustion in a year. More frequent wildfires also threaten to increase carbon emissions. And as habitats are made untenable, what will be the effect on biodiversity?

At the same time, drought models need to be improved. They successfully chart the hydrological changes seen in the US Southwest and the drying seen at the global level<sup>7</sup>, but regional predictions can be disturbingly variable. Some models forecast an increase in precipitation for East Africa, whereas others correctly predicted in 2010 that warming of the Indian Ocean would lead to drought in the region, such as this year's devastating drought in Somalia. The models need higher resolution and a better understanding of precipitation, sea surface temperature and the effects of vegetation.

Human adaptation to prolonged,

extreme drought is difficult or impossible. Historically, the primary adaptation to dust-bowlification has been abandonment; the very word 'desert' comes from the Latin *desertum* for 'an abandoned place'. During the relatively short-lived US Dust-Bowl era, hundreds of thousands of families fled the region. We need to plan how the world will deal with drought-spurred migrations (see page 447) and steadily growing areas of non-arable land in the heart of densely populated countries and global bread-baskets. Feeding some 9 billion people by mid-century in the face of a rapidly worsening climate may well be the greatest challenge the human race has ever faced.

These predictions are not worst-case scenarios: they assume business-as-usual greenhouse-gas emissions. We can hope that the models are too pessimistic, but some changes, such as the expansion of the subtropics, already seem to be occurring faster than models have projected<sup>10</sup>. We clearly need to pursue the most aggressive greenhouse-gas mitigation policies promptly, and put dust-bowlification atop the world agenda. ■

**Joseph Romm** is a physicist who edits the blog *ClimateProgress.org* for the Center for American Progress Action Fund, Washington DC, USA.  
e-mail: [jromm@americanprogress.org](mailto:jromm@americanprogress.org)

1. Rind, D. *et al.* *J. Geophys. Res.* **95**, 9983–10004 (1990).
2. Barnett, T. P. *Science* **319**, 1080–1083 (2008).
3. Stewart, I. T. *et al.* *J. Climate* **18**, 1136–1154 (2005).
4. Woodhouse, C. A. *et al.* *Proc. Natl Acad. Sci. USA* **107**, 21283–21288 (2010).
5. *Global Climate Change Impacts in the U.S.* (USGCRP/Cambridge Univ. Press, 2009).
6. Seager, R. *et al.* *Science* **316**, 1181–1184 (2007).
7. Dai, A. *J. Geophys. Res.* **116**, D12115 (2011).
8. Dai, A. *Climate Change* **2**, 45–65 (2011).
9. Solomon, S. *et al.* *Proc. Natl Acad. Sci. USA* **106**, 1704–1709 (2009).
10. Seidel, D. J. *et al.* *Nature Geosci.* **1**, 21–24 (2008).



J. AHERAM/AHERAM.COM

Despite multiple conflicts and two world wars in the past century, societal evolution has led to an overall decline in violence and death.

## PSYCHOLOGY

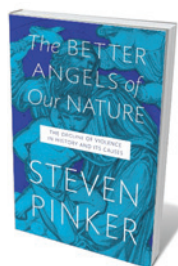
# A farewell to arms

Martin Daly explores Steven Pinker's treatise on the taming of human aggression.

**W**ars and genocides, murder and mayhem — violent victimization seems to be rising inexorably. But the massive coverage of these horrific phenomena masks an important truth. On average, our chances of being assaulted or killed have been falling for centuries. In *The Better Angels of Our Nature*, psychologist Steven Pinker reviews the evidence for this stunning historical trend, and tries to explain it.

In what is arguably his most ambitious work yet, Pinker includes figures showing declining rates of homicide, warfare, acts of terrorism, child abuse and other forms of violence over various timescales. But there is more here than statistics. Pinker's narrative moves from prehistory, through the social and intellectual revolutions of the eighteenth and nineteenth centuries, to current findings on mind, brain and behaviour. Citing the insights and scholarship of not just the usual gang of psychologists, neuroscientists and evolutionary biologists, but also of historians, philosophers and every sort of social scientist, he concludes that societal evolution has reduced the incentives to commit violence and changed modern sensibilities.

The intellectual hero of Pinker's story — a



**The Better Angels of Our Nature: Why Violence Has Declined/The Decline of Violence in History and Its Causes**

STEVEN PINKER  
Viking/Allen Lane:  
2011. 832 pp.  
\$40/£30

*Leviathan*) that could monopolize legitimate violence and arbitrate disputes, reducing the need for private retribution. The other was the rise of 'gentle commerce', whereby mutual gains from trade created a common purpose.

After introducing and defending Elias's ideas, Pinker applies them to history. Reviewing the 'humanitarian revolution', he

man he calls "the most important thinker you have never heard of" — is the German-born sociologist Norbert Elias. In the late 1930s, he proposed that recent centuries have witnessed a 'civilizing process' in the form of a growing regard for others' interests and improved self-control. According to Elias, the civilizing process had two causes. One was the consolidation of governments (as described in Thomas Hobbes's 1651 book

describes how, in Europe and elsewhere, the xenophobia that was once ubiquitous became untenable over several centuries. Torture, execution at whim and slavery also moved from the mainstream to the marginal.

Pinker then makes the case that warfare has long been in decline — and may be facing extinction. Scholarly analysis of armed conflicts and their death tolls apparently demonstrates that a bias towards recency has blinded us to this startling truth: that both the incidence of war and the death rate it imposes have been shrinking. Even in the twentieth century, with its two world wars, these numbers were lower than in previous centuries. And they have kept falling. Pinker chronicles the 'rights revolutions' of the twentieth century: struggles for civil rights, women's rights, children's rights, gay rights and animal rights, with thought-provoking discussion of the rapidly changing sensibilities that accompanied them. His skill in mixing quantitative analysis with illustrative examples, apt quotations and the occasional joke makes these chapters page-turners.

Pinker then turns his attention to the links between the history of violence and his view — familiar from his previous books — that the evolved human psyche is a bundle of special-purpose 'faculties', including certain "inner demons" and "better angels". He reviews what neuroscience has to say about aggression and empathy, and what social psychologists have discovered about the elicitors of sympathy, punitiveness and other mental states. He concludes with a surprising, forceful argument for a

➔ [NATURE.COM](http://NATURE.COM)

For Steven Pinker's views on declining violence:  
[go.nature.com/nbrlsp](http://go.nature.com/nbrlsp)

historical rise in human reasoning ability, which he believes may provide a counterweight to parochialism and intergroup hostility.

Any thoughtful reader of this wide-ranging treatise will find nits to pick. I am not persuaded that because 1960s counterculture glorified selfish impulsiveness ('just do it'), it was responsible for a small increase in the US homicide rate. The counterculture was largely about forsaking violence ('make love, not war'). Neither am I persuaded by Pinker's explanations for the decline in US homicide in the 1990s; in my view, he too hastily dismisses the possible relevance of demographic changes, and too credulously accepts that increased police presence and incarceration were important.

Pinker's biggest slip, in my view, concerns the relevance of income inequality, which has been the most successful predictor of variability in homicide rates between different places. Pinker gives it one brief paragraph, waving it off on the grounds that the standard index of income inequality was going up during the 1990s in the United States while crime rates were falling, and was at a nadir in 1968 when crime was "soaring". The trouble with this argument is that there is no reason to expect simultaneous short-term vicissitudes of income inequality and homicide; any effect of the former on the latter is surely mediated by people's cumulative experiences over their lifetimes. And it is ironic that despite Pinker's dismissal, the big historical story he tells — stressing the decline of despotism and marauding and the rise of democratic governments — is itself a tale of decreasing inequality.

Pinker closes with a rousing defence of modernity. Ultimately, his explanation for the decline of violence is Elias's — that the synergistic impacts of *Leviathan* and gains from trade have created a civilizing process that has diminished the utility of violence and, hence, its appeal. But he elaborates on this with an engaging game-theoretical twist (the "Pacifist's Dilemma"), and more-up-to-date psychology than Elias would have been able to muster. *The Better Angels of Our Nature* is a lively, fascinating read and a remarkable scholarly achievement that deserves to be studied and debated by many social scientists, concerned citizens and policy-makers. ■

**Martin Daly** is professor of psychology in the Department of Psychology, Neuroscience and Behaviour, McMaster University, Ontario L8S 4K1, Canada. e-mail: daly@mcmaster.ca



A biology graduate's back carries a reminder that DNA gave rise to all the biodiversity on Earth.

#### SOCIOLOGY

## The illustrated scientist

Margo DeMello is fascinated by the evocative tattoo culture among different 'tribes' of scientists.

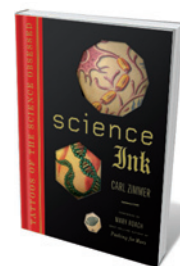
Tattoos were taboo until recently in the West — seen by most as the barbaric practice of marginalized or under-world groups. Now, tattooing is undergoing a renaissance. Almost mainstream in Europe and North America, tattoos are becoming ever more artistically sophisticated and personally meaningful.

Carl Zimmer's beautiful new book, *Science Ink*, focuses on tattoo culture among scientists, both amateur and professional. Zimmer, himself a tattoo-free science writer, began asking researchers to send photographs of their science-related tattoos to The Loom, his blog for *Discover Magazine*, in August 2007. These, and the stories behind them, evolved into *Science Ink*.

The book is broadly organized by discipline, featuring photos of tattoos themed to each — astronomy, chemistry, evolution, natural history, neuroscience and palaeontology. The scientists are using their body art to mark their standing as members of these 'tribes': so you see stars on astronomers, bacteria on biochemists, insects on entomologists and equations and symbols on mathematicians. And

there are molecules of every type, including pages of double helices.

Some designs are iconic, such as  $E=mc^2$ ; or personal, like the chemist's tattoo of the molecular structure of phenobarbital, a drug he gives to his cat to control its seizures. One Loom reader sent in a sequence of zeroes and ones — the name of his daughter Lain in binary code. Some of



**Science Ink:**  
**Tattoos of the**  
**Science Obsessed**

CARL ZIMMER  
Sterling: 2011. 288 pp.  
\$24.95, £16.99

the tattoos are simple line drawings. Many are colourful and stunningly detailed — such as the elaborate picture on a mathematician's back of a microscope and the usually hidden world it reveals.

These decorated scientists join a tradition that is both venerable and near universal. The earliest evidence for tattooing dates back to Neolithic Eurasia. From there it probably spread from the Middle East to the Pacific

Islands, and later to the Americas, by way of India, China and Japan. By 3,000 years ago it was found almost everywhere, and today remains rare only in sub-Saharan Africa.

As permanent body art, similar to scarification, tattoos typically marked permanent or semi-permanent aspects of social position, such as rank or marital status. Today, they still serve this purpose, among others. As I wrote in *Bodies of Inscription* (Duke University Press, 2005), when the middle classes began getting tattoos, they also began to create “tattoo narratives”: stories relating why they got the tattoo, how long they had thought about it, the genesis of the design and its meaning, the tattooing experience and what the tattoo means to them now.

For professionals, these narratives are particularly important. As trailblazers in their class, they need to create new meanings for their tattoos; underworld or working-class narratives are not relevant to them. New narratives are important for personal as well as social and ‘tribal’ reasons — the scientists don’t want their choices to seem random or impulsive.

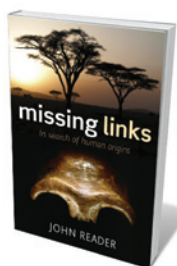
Many of the scientists’ designs are not easily understandable without knowing the story behind them. For example, the tattoo that inspired Zimmer’s quest was a double helix acquired by one of his friends, a neurobiologist. But it isn’t just any DNA: it also spells out the name of that friend’s wife. Another couple featured in the book have matching tattoos of chromosomes splitting during meiosis; those with no basic understanding of biology would have a hard time grasping the literal or metaphorical meaning of their squiggles without a narrative.

*Science Ink* is packed with fascinating stories. One of the most moving is Abigail’s. A chemistry student, she sent in a photo of her tattoo — the word ‘entropy’ inked on her back. A few months later, her mother sent Zimmer a note saying that Abigail had died in a car accident and that she was getting her daughter’s tattoo replicated on her own body. That blog post and the comments it generated became a memorial for Abigail, and eventually led to a posting by a woman whose mother had received Abigail’s lungs after her death.

We call tattoos permanent, but they last only as long as the body that wears them survives. Abigail’s tattoo has a life beyond her own: the design now adorns the headstone marking her grave. And it is there in the pages of *Science Ink* — one of many signs of an enduring fervour for science, and a new chapter in the age-old history of body art. ■

**Margo DeMello** is a cultural anthropologist and author of *Bodies of Inscription: A Cultural History of the Modern Tattoo Community*.  
e-mail: [margo@animalsandsociety.org](mailto:margo@animalsandsociety.org)

## Books in brief



### Missing Links: In Search of Human Origins

John Reader OXFORD UNIVERSITY PRESS 350 pp. £25 (2011)

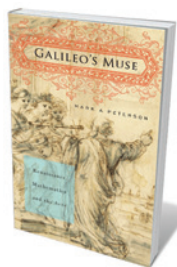
The cast of ancient superstars in palaeoanthropologist John Reader’s book has grown significantly in the 30 years since the first edition. Neanderthal Man, Lucy and other early hominin fossils are joined by finds from *Homo floresiensis* to *Ardipithecus* in a stunningly illustrated update. Powered by enthusiasm and peppered with controversy, the search for human origins is laid out clearly and succinctly, from the first fossils and Victorian revelations, to frauds such as Piltdown Man and triumphs such as the ‘world’s oldest child’: the *Australopithecus afarensis* fossil unearthed in Ethiopia and called Selam (‘peace’).



### American Madness: The Rise and Fall of Dementia Praecox

Richard Noll HARVARD UNIVERSITY PRESS 390 pp. £33.95 (2011)

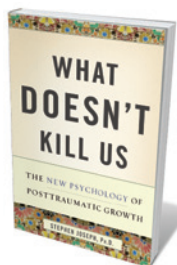
Between 1895 and the 1930s, tens of thousands of Americans were diagnosed with dementia praecox — an ‘incurable’ psychosis described by German psychiatrist Emil Kraepelin. The diagnoses then petered out. Psychologist Richard Noll traces the trajectory of this near-forgotten disorder, showing how it became the first specified disease of psychiatry, legitimizing that field’s place in medicine. Noll also shows how the debates today around the successor to dementia praecox, schizophrenia, are leading to a trend in psychiatry towards diagnoses that could fit better with genetics.



### Galileo's Muse: Renaissance Mathematics and the Arts

Mark A. Peterson HARVARD UNIVERSITY PRESS 336 pp. £21.95 (2011)

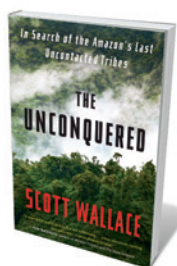
The great scientist Galileo Galilei was also a gifted draftsman and accomplished musician, steeped in Renaissance poetry. But art was no side interest for Galileo, physicist Mark Peterson claims. The mathematical inspiration for his findings, such as four of Jupiter’s moons, was fished from the humanist stream then flowing so powerfully in Italy. So it was Dante’s *Inferno*, Filippo Brunelleschi’s great domes and artist-innovators from Piero della Francesca to Leonardo da Vinci, not the medieval tag ends of science, that inspired Galileo and ignited the Enlightenment, Peterson argues.



### What Doesn't Kill Us: The New Psychology of Posttraumatic Growth

Stephen Joseph BASIC BOOKS 288 pp. \$26.99 (2011)

Tsunamis, assault, near-death accidents: such experiences are popularly imagined to scar victims ‘for life’ and leave them in thrall to post-traumatic stress disorder. After two decades of research, positive psychologist Stephen Joseph argues that, for many, these traumas can become an “engine for transformation”. Backed by case studies, he covers trauma’s emotional toll, the underlying biology, the realities of resilience and the array of therapies on offer, such as trauma-focused cognitive behaviour therapy. This is a thorough and common-sense look at the psychology of survival.



### The Unconquered: In Search of the Amazon's Last Uncontacted Tribes

Scott Wallace CROWN 512 pp. \$26 (2011)

Conquering civilizations have ebbed and flowed through Latin America, but uncontacted tribes such as the *flecheiros* (or Arrow People) still survive deep in the Amazon rainforest. Now their home and culture are threatened by deforestation, epidemics and marginalization. Journalist Scott Wallace takes us on a journey through a warzone where irreplaceable habitats and the knowledge of traditional peoples are the casualties.

## BIOTECHNOLOGY

# DNA dollars

**Linnaea Ostroff** examines a history of Genentech, the US company that first made biology a business.

As the mysteries and mechanics of DNA were being revealed, it was unclear whether the molecule would be used for good or evil. Debates raged: utopian fantasies of ending disease and famine competed with fears of mutated life forms running amok. A suitably startling, if less popcorn-worthy, event occurred in October 1980, when the promise of DNA modification raised US\$35 million in a landmark initial public offering (IPO), which saw the fastest stock-price rise in the market's history.

The record-breaking IPO was that of Genentech, a small company based in San Francisco, California, whose plan was to produce drugs using recombinant DNA technology. This was the first commercial manipulation of DNA and the first sale of biological science as a commodity in its own right. The biotech industry was born. Genentech's unique corporate structure, which blurred the boundary between academia and industry, was swiftly imitated. The sometimes uncomfortable entanglement of publicly funded basic research with private business enterprise persists to this day.

*Genentech* by science historian Sally Smith Hughes gives a detailed account of the founding and early years of the company. Much of the material in the book comes from oral histories collected by Hughes, along with written archival material. Hughes's book is not, however, a journalistic analysis of a unique and important company: it is an account of the key players, as told to a sincere admirer.

Nevertheless, Genentech's achievements in science, medicine and business were momentous. One of the company's co-founders, Herbert Boyer, a molecular biologist at the University of California, San Francisco, was at the time a leader in the development of recombinant DNA technology. Boyer and others had recently discovered a means of reorganizing (recombining) the sequence of DNA molecules, and were pursuing a method to use this engineered DNA to generate proteins. This had

profound implications for drug production and development.

Whereas most drugs had been discovered by large-scale screening of synthetic chemicals, a handful, such as insulin, were natural proteins whose production in the body was impaired in diseases such as diabetes. Proteins have exceptionally complex structures, and it is still too difficult to routinely synthesize them from scratch. Therapeutic proteins were at the time sourced from animals' organs and human cadavers, making their supply and safety unreliable. In theory, recombinant DNA could provide a safe, consistent source of this class of therapeutics.

Boyer's group was working on a way to coax bacterial cells to produce therapeutic proteins from recombinant DNA. More importantly, recombinant DNA presented a means of designing drugs using the biological mechanisms of a particular disease, which seemed to be an obvious advance over the pharmaceutical industry's random screening procedures. Hughes does not, however, touch on any of this, leaving the reader to wonder why recombinant DNA is viewed as so useful.

The reasons the IPO was so successful, and why that success was so shocking, are also underdeveloped in the book. At the time it went public, Genentech had the intention of making pharmaceuticals but had no actual drugs in the pipeline. What it did have was a contract with Eli Lilly, the largest producer of synthetic insulin. The contract was the first of its kind: Eli Lilly was not paying Genentech to produce insulin, nor licensing a method to do so, but was paying it to do the basic scientific research needed to develop a method. Never before had an independent group of scientists contracted with a for-profit organization to make basic scientific discoveries, nor had a publicly traded



**Genentech: The Beginnings of Biotech**  
SALLY SMITH HUGHES  
University of Chicago  
Press: 2011. 232 pp.  
\$25, £16

company offered research as its sole source of revenue.

A patent on recombinant DNA techniques was granted in 1980 to Stanford University, California, and to the University of California, where Boyer and his colleagues had developed the technology. The assurance of intellectual-property protection for genetic-engineering methods and products encouraged the explosion of the biotechnology sector, as academic researchers began to independently commercialize their findings. The now commonplace practice of scientists maintaining ties to both universities and their own associated companies, along with the conflicts it creates, comes directly from Genentech's initial arrangement.

Although Genentech's business model was groundbreaking in its mechanics, the long-term strategy was a classic risk. Genentech's insulin was intended to be the Gutenberg Bible of recombinant DNA technology — an established product made in a new way with a guaranteed market. Yet the route between basic knowledge of a disease process and an effective therapy is punishing, and many subsequent designer drugs generated using the method proved not to be viable.

Rational drug design has not overtaken traditional drug-discovery approaches, and biotechnology development is shifting back to large pharmaceutical companies, which can hedge risk internally — although the future of drug discovery is a legitimate concern. Genentech itself is now wholly owned by Swiss pharmaceutical giant Roche.

The scant objectivity, the somewhat plodding chronology of unfolding events and the sparse explanations of technical terminology in Hughes's account aside, Genentech's story remains a compelling one. It neatly reveals the divergent challenges of basic science, medical science and business, and despite its novelty, the tale illustrates several enduring principles of science and markets.

In shifting genetic-engineering research from academia to industry, Genentech and the industry it founded accelerated the development and distribution of medically and agriculturally valuable products. It triggered practical decisions on policy and regulation, while effectively sidestepping philosophical and ethical questions about the uses of DNA: the market would decide what DNA should be used for. Genentech's business model shunted private money directly into basic research, drew investors into basic science and academic scientists into business. Even as the industry reorganizes, these relationships remain. ■

**Linnaea Ostroff** is a researcher at the Center for Neural Science, New York University, New York 10003, USA.  
e-mail: lostroff@nyu.edu

➔ **NATURE.COM**  
For a review of the US biotech debate:  
[go.nature.com/yv1ffo](http://go.nature.com/yv1ffo)



## Q&A Persi Diaconis

# The mathematician

*Mathematician Persi Diaconis of Stanford University in California ran away from home in his teens to perform card tricks. As he publishes a book on the mathematics of magic, co-authored with juggler and fellow mathematician Ron Graham, he explains what makes a good trick.*

### Which came first for you, magic or maths?

Magic came first. When I was five, I found a magic book in the attic and started doing shows. I was a terrible magician but the other kids liked it. In high school I had a good geometry teacher but had no interest in mathematics and never did homework. When I was 13, I met Alex Elmsley, a soft-spoken British computer engineer and magician, at a magic shop in New York City. He showed me that eight perfect shuffles would put a deck back in its original order. His ingenious method for moving the top card to a desired position within the deck was my introduction to binary numbers.

### How did you pursue magic?

The sleight-of-hand artist Dai Vernon invited me to Delaware to do some magic shows when I was 14 — and I never went back home. We found crooked gamblers to learn their techniques. There are often tricky probability questions involved. I didn't have much probabilistic intuition and made every boneheaded mistake you could make. During ten years on the road, I learned the hard way.

### Can you describe one of the best tricks you learned?

In 1916 the pioneering US magician Charles Thornton Jordan advertised a trick called

### Magical Mathematics: The Mathematical Ideas that Animate Great Magic Tricks

PERSI DIACONIS AND RON GRAHAM  
Princeton University Press: 2011. 258 pp.  
\$29.95, £20.95

returned the deck to him, he would name your card. It is a wonderful trick that fooled everybody and didn't seem mathematical. It works because the deck is arranged in a special order. When you shuffle once, the deck is split into two alternating sequences (or 'chains'). Two shuffles makes four interlocking chains, and three shuffles makes eight chains. When he got the deck back, he would play a sort of solitaire to isolate the one card that was not in any of these chains. That was the chosen card.

### How did you come to study shuffling theory?

Jordan's trick led me to ask how many times you have to shuffle a deck of cards to mix them up properly. People often ask why this problem can't be solved by brute force with a computer. But a deck of cards can be arranged in almost as many ways as there are atoms in the Universe. All the computers

'Long-Distance Mind Reading'. He would mail you an ordinary deck of cards, ask you to cut and shuffle twice, then draw a card and restore it to the middle of the deck. After you shuffled again and

in the world couldn't run through all the arrangements. We used probability theory, combinatorics and group theory to prove that you need seven ordinary riffle shuffles to mix up a deck randomly. Now we're using the mathematics of shuffling to study turbulence in fluid dynamics, which has many industrial applications, such as determining how long a vat of cookie dough must be blended to ensure that all of the ingredients are mixed.

### Are there any other practical applications of card tricks?

Jordan invented a method of ordering a deck such that the pattern of reds and blacks in a series would code for a unique set of cards, allowing him to divine cards on very little information. Such an arrangement is known as a De Bruijn sequence among mathematicians. These sequences have lots of practical applications. They are used to scramble mobile-phone signals, reassemble snippets of DNA and allow a digital pen to identify its position on special paper.

### Could science explain what makes a good magic trick?

The psychology of deception is a serious subject, particularly among spy agencies, but I have never seen a convincing study of it. Some magic tricks are viscerally moving and shocking, others are painfully boring. Dai Vernon said that good magic has a way of "ingeniously leading the mind to defeat its own logic". That sort of thing is not so scientific. Magic is a theatrical experience.

### How are secrets treated among magicians?

It is a strange tension. In magic there are still many secrets. I'm famous for keeping them. If someone shows me something, it stays with me forever. When a student asked me recently to talk about card tricks, I declined. But Wikipedia and YouTube are changing things. Someone can find the magician's secret on their phone during a performance. Maybe this will make magicians more inventive, or make people more appreciative.

### What kind of maths do you prefer?

I enjoy learning new things. When you start in a new field you have to ask dumb questions. I often say I'm paid for my ability to tolerate feeling stupid. I also like problems that touch the real world, and that you can explain simply, like flipping a coin or spinning a roulette wheel. When I develop a big piece of theory, I feel like I'm slumming: the real discoveries are in the examples. This spring, I will teach a course in the mathematics of magic tricks, in an effort to twist young minds in the right direction. ■

INTERVIEW BY JASCHA HOFFMAN

# Correspondence

## Iraq's unseen burden of wounded civilians

The burden of war-related mental disorders is well documented among US veterans (*Nature* **477**, 390–393; 2011), but not among civilians in Iraq. This oversight must be rectified so that adequate medical support can be provided to the Iraqi people.

US combat troops will soon depart Iraq, leaving Iraqis to cope with the consequences of the 2003 invasion. Although the number of violent deaths is falling, civilians have been killed almost every day this year, most of them in coordinated bomb attacks. Roadside blasts cause long-term disabilities and societal effects among injured civilians. However, these have been largely neglected by the media and no systematic surveillance has been undertaken.

Despite Iraq's damaged health-care system, primary repairs of many injuries are being carried out with acceptable results. But some patients require advanced procedures that are not available in the country. For international aid organizations that are unable to work safely inside Iraq, one operational model is to treat patients outside the country.

Médecins Sans Frontières (MSF; also known as Doctors Without Borders) started a surgical programme in Amman, Jordan, in 2006 to provide functional reconstructive surgery and psychosocial support for Iraqi civilians. Although MSF has managed progressively to increase the capacity of its programme, it can still accommodate only a small fraction of the Iraqi civilians in desperate need of advanced surgical care.

Development of standardized data-collection tools would greatly improve future monitoring of mental health and of explosion injuries. Culturally adequate interventions and tailored support networks are needed to alleviate the long-term physical and psychological

repercussions of exposure to war-related trauma. Above all, efforts must concentrate on improving security for civilians.

**Gilles Guerrier, Emmanuel Baron** *Epicentre, Paris, France.*  
[guerriergilles@gmail.com](mailto:guerriergilles@gmail.com)  
**Rasheed Fakri** *Médecins Sans Frontières, Amman, Jordan.*  
**Isabelle Mouniaman** *Médecins Sans Frontières, Paris, France.*

## Brazil's forest code puts wetlands at risk

Brazil's revisions to its Forest Code threaten not only the Amazon rainforest but also its wetlands (*Nature* **476**, 259–260; 2011). Many Brazilian flood plains extend into neighbouring countries, so they could also be affected.

Seasonal rainfall causes the levels of most Brazilian rivers to fluctuate. Flood plains reach widths of tens of metres along small streams and tens of kilometres along large rivers, and up to 90% of these dry up during periods of low rainfall.

These wetlands provide the environment and humans with important services, such as water storage, discharge buffering, water clearing, sediment retention, recharging of the groundwater level, local and regional climate regulation, and maintenance of a large biodiversity. Some provide homes and livelihoods for traditional human populations as they harbour important fish stocks and can also be managed for low-density cattle ranching and timber production.

Neither the old nor the new version of the Forest Code specifically mentions wetlands. The old code protects forests along streams and rivers, according to the river's width and maximum water level, thereby integrating and protecting the wetland areas. The new code protects areas only to a poorly defined "regular" water level, opening up opportunities for the destruction of high-lying

wetland areas. This will damage the integrity of the remaining low-lying areas, along with most of their benefits for humans and the environment.

The Intergovernmental Panel on Climate Change predicts that large parts of Brazil will experience heavier rainy seasons and more severe dry periods interspersed with heavy rainfall. Buffering by intact wetlands will be increasingly important as water availability and distribution become the limiting factors for agricultural development and the well-being of rural and urban populations.

**Paulo Teixeira de Sousa Jr** *National Institute for Science and Technology in Wetlands (INAU), Cuiabá, Mato Grosso, Brazil.*  
[pauloteixeiradesousa@gmail.com](mailto:pauloteixeiradesousa@gmail.com)  
**Maria Teresa Fernandez Piedade** *National Institute of Amazonian Research (INPA), Manaus, Amazonas, Brazil.*  
**Ennio Candotti** *Museum of Amazônia, Manaus, Amazonas, Brazil.*

## Small colleges aided by research networks

As faculty members of primarily undergraduate institutions (PUIs), we have successfully developed research programmes with our students despite funding and collaborator limitations (*Nature* **477**, 239–241; 2011). Long-term research at PUIs is challenging, but feasible.

To improve our research productivity, we have set up collaborative research networks with other PUIs. Through our Ecological Research as Education Network (EREN, comprising 72 PUIs), we are helping each other to develop grant proposals, research protocols, data sets and manuscripts, and are teaching students through multi-institutional, collaborative research (D. R. Bowne *et al.* *BioScience* **61**, 386–392; 2011).

Internal research funding, including start-up packages, varies widely among institutions.

In a survey of 50 ecology faculty members from PUIs at the Ecological Society of America's 2011 annual meeting, 30% of respondents reported no internal support for research equipment and supplies, and 51% said there was none for student-researcher stipends. Inter-institutional networks enable faculty members to share research resources at minimal cost.

PUI faculty members are creative in seeking research funding. For example, the same survey revealed that 94 funding sources had been successfully accessed, including US government agencies.

PUI faculty members often collect long-term data during undergraduate courses and in independent research, although better coordination is needed to enhance the scientific and educational impact of this work. Members of EREN have shared research protocols with many institutions to answer continental-scale questions.

**Erin S. Lindquist** *Meredith College, Raleigh, North Carolina, USA.*  
[erinlind@meredith.edu](mailto:erinlind@meredith.edu)  
**Laurel J. Anderson** *Ohio Wesleyan University, Delaware, Ohio, USA.*

**Jeffrey A. Simmons** *Mount St. Mary's University, Emmitsburg, Maryland, USA.*

## Aboriginal people agreed to DNA study

As research manager of the Goldfields Land and Sea Council (GLSC), I was involved in your discussion of Aboriginal genome research (*Nature* **477**, 522–523; 2011) and would like to make it clear that the decision to allow analysis of the 90-year-old hair sample was made by the duly mandated people. The decision took proper account of ethical research practices and of the rights of Aboriginal people to safeguard their cultural heritage.

The GLSC is the representative body for the Aboriginal people

in the region where the sample was obtained, and is recognized under the Native Title Act 1993. The directors are elected by GLSC members, and membership is open to all Aboriginal residents of the region. In granting their permission for the research, the board exercised properly defined moral, cultural and legal authority to speak on behalf of the Aboriginal people there.

Most research — be it sociological, historical or genetic, or even political polling — extrapolates from a sample to draw conclusions. Participants are rarely expected to seek consent from their entire group before giving up information.

Because the hair sample was almost certainly given to British ethnologist Alfred Cort Haddon voluntarily in the early 1920s, this example of an informal exchange between an Aboriginal person and a researcher does not provide a model for all such exchanges in the future. These should be underpinned by a

standard indicating that free, prior and informed consent was sought from the proper people.

**Craig Muller** *Goldfields Land and Sea Council, Perth, Western Australia, Australia.*  
[craig.muller@glc.com.au](mailto:craig.muller@glc.com.au)

## Give more priority to phosphorus studies

I agree with James Elser and Elena Bennett that we should recycle phosphorus (*Nature* **478**, 29–31; 2011). However, there are situations in which the natural recycling of phosphorus is not ecologically desirable.

As the authors note, excess phosphorus in water bodies can feed algal blooms and create anoxic zones. What is less well known is that these waters can become permanent dead zones, stuck in an oxygen-deprived, nutrient-rich state. This happens when the algae die, sink and are decomposed by anaerobic bacteria that need only limited

amounts of phosphorus. Most of the algal phosphorus is released back into the water to feed further blooms. In the Baltic Sea, for example, reductions in phosphorus pollution from rivers have not yet led to ecosystem recovery because of this effect.

We have disturbingly little insight into major phosphorus fluxes in the marine realm. This is the legacy of decades of research priority being given to the microbial complexities of the nitrogen cycle over the methodologically challenging investigation of phosphorus cycling.

**Caroline P. Slomp** *Utrecht University, the Netherlands.*  
[c.p.slomp@uu.nl](mailto:c.p.slomp@uu.nl)

## Boost resilience to tackle mental illness

An economically efficient way of tackling the enormous social and economic costs of

mental ill health (*Nature* **477**, 132 and **478**, 15; 2011) would be to boost 'resilience' to mood disorders.

Mentally healthy individuals often show a positive affective bias because their processing of negative information is inhibited. This effect, possibly mediated by the neuromodulator serotonin, promotes resilience by dampening the stress associated with negative life experiences.

Devising ways to promote such resilience in healthy individuals could help to prevent chronic stress-related brain disorders, saving huge amounts of money and heartache every year.

**Oliver J. Robinson** *National Institute of Mental Health, Bethesda, Maryland, USA.*  
[robinsonoj@mail.nih.gov](mailto:robinsonoj@mail.nih.gov)

**Disclaimer:** Views presented in this Correspondence are solely those of the author and do not necessarily represent the views of the US federal government.



## NATURE'S READERS COMMENT ONLINE

*Selected responses to 'Fund people not projects' by John P. A. Ioannidis (Nature 477, 529–531; 2011).*

### Yiding Zhao says:

'Fund people not projects' was once the model used in China, but major international journals frowned on it because it risked creating *Xue ba* (scientific autocracy that suppresses others' ideas). So we worked hard to adapt the grant-based model. Now you are telling us the grant-based model is worse?

[yidingzhao@pku.edu.cn](mailto:yidingzhao@pku.edu.cn)

### Ken Whitmire says:

A way to fund people rather than projects would be to allocate money directly to individual graduate students and postdocs through fellowships, instead of funnelling it through a principal investigator's grant. This redirection wouldn't cost the system any more money, and it would make it clear to non-scientists that fellowships are funding the training of a highly skilled technical workforce, as well as helping a research enterprise. Students would have more independence in choosing an adviser and advisers would

be under less pressure to raise huge sums of money to support an active research group.  
[whitmir@rice.edu](mailto:whitmir@rice.edu)

### Sander Heinsalu says:

Funding models have trade-offs. Detailed checks create bureaucracy, but avoid misuse of money. Specific goals limit creativity, but avoid funding less useful projects. I would like to see scientific evidence on which funding models generate better output — although the best definition of output is also debatable.

[sanderheinsalu@hotmail.com](mailto:sanderheinsalu@hotmail.com)

### Adrian Barnett says:

One option is to fund projects retrospectively, with money being handed out for work delivered (including papers, policy changes, improvements in health), rather than for promises made in grant applications. Most current grant systems are heavily biased towards senior staff, but this scheme would work irrespective of applicants'

status. The process would be less burdensome for researchers because it would involve gathering their existing evidence and costs.  
[a.barnett@qut.edu.au](mailto:a.barnett@qut.edu.au)

### Craig Macfarlane says:

The funding model is largely irrelevant — what really matters is the amount of money. Whatever the system is, scientists are smart enough to learn to play it, and it will be dominated by established players who are closest to its centre. The only solutions are for developed countries to increase public funding for research to ensure that it is not just the heavyweights who receive grants, and to support more public-good research. Anything else is just fiddling at the margins. Until that happens (when hell freezes over perhaps), get used to things the way they are.  
[fisheye@iinet.net.au](mailto:fisheye@iinet.net.au)

**To join this debate, go to**  
<http://go.nature.com/ciahh5>.

# Ralph Steinman

## (1943–2011)

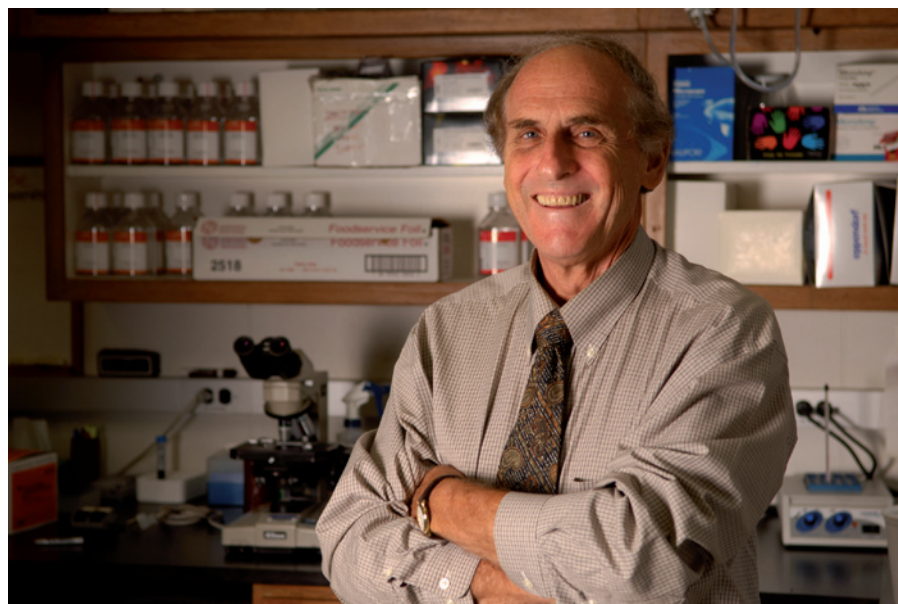
Immunologist and cheerleader for dendritic-cell biology.

Ralph Steinman changed the world of immunology when he discovered dendritic cells, but it took the field a long time to recognize the importance of his discovery. The idea that a new type of immune cell could be found in 1973 — in the era of molecular cell biology — simply by looking down a microscope seemed far-fetched. The early criticism was relentless. Steinman's road to the Nobel Prize in Physiology or Medicine — awarded (unusually) just days after his death — was full of obstacles. That the journey was even possible was down to his forceful personality, energy and focus.

Steinman was born in 1943 in Sherbrook, Quebec, Canada, as the second son in a family of Jewish immigrants originating in Moldavia and Poland. The Steinmans owned Mozart's, a general store selling everything from appliances to clothing. His parents wanted him to study religion and take over the family business, but summers working in the store reinforced Steinman's desire to do something else. His love of science led him to McGill University in Montreal, Canada, then to Harvard Medical School in Boston, Massachusetts, where in the late 1960s he heard lectures by Kurt Bloch on the initiation of immunity and studied Peter Medawar's work on tolerance and Frank McFarlane Burnet's ideas on clonal selection.

Although Steinman did a residency in internal medicine at Massachusetts General Hospital in Boston, he was drawn to basic research. In 1970, he joined Zanvil Cohn and James Hirsch's laboratory at Rockefeller University in New York to work on the initiation of immune responses. Cohn and Hirsch were focusing on macrophages, but Steinman was also influenced by his campus neighbours: Christian de Duve, George Palade, Philip Siekevitz, David Sabatini and Günter Blobel were inventing modern cell biology a few floors above. Steinman soon characterized pathways for the engulfing of molecules by cells (endocytosis) and, with Cohn, proposed the involvement of membrane recycling in this process.

**“As a basic scientist he recognized the enormous challenge of taking a discovery from the laboratory to the patient.”**



ROCKEFELLER UNIV.

At that time, in the early 1970s, immunologists were developing culture systems to help study the cellular basis of immunity. An early finding was that, in addition to B and T lymphocytes, antibody responses required another type of cell, dubbed an ‘accessory cell’. The mysterious accessory cells stuck to glass, so Steinman — inspired by his cell-biology colleagues’ emphasis on microscopy — decided to examine glass-adherent spleen cells using phase contrast, live imaging and electron-microscopy techniques.

What he saw under the lens was a new type of immune cell that had branching, rapidly changing projections. Three experiments (on two of which, one of us, MCN, was involved) then convinced Steinman that these ‘dendritic’ cells were the missing accessory cells: they could induce T-cell division and initiate killer-T-cell responses to antigens, and were biochemically distinct from macrophages.

For decades, Steinman was the consensus leader of, and the most enthusiastic cheerleader for, the field of dendritic-cell biology. He brought in numerous scientists from other fields, and loved to collaborate. With Wesley van Voorhis, he showed that dendritic cells exist in the blood of humans. With Kayo Inaba, he established, among other things, that when loaded with antigen, dendritic cells could induce anti-tumour immunity in mice. With Gerald Schuler, he saw that dendritic cells

could be activated by pathogens to initiate immunity.

Steinman was generous and much loved by his colleagues. Even when he was terribly ill, he spent the little time he had left ensuring that his students and fellows would land on their feet after he passed.

He was passionate about his work as editor of the *Journal of Experimental Medicine*, a responsibility that he enjoyed for more than 40 years. His focus on publishing outstanding science had a profound influence on his field.

And he was equally passionate about making the leap from the bench to the bedside. In recent years, he tried to use dendritic cells to develop vaccines. As a basic scientist he recognized the enormous challenge of taking a discovery from the laboratory to the patient. But he relished the task because of its importance in treating infectious diseases and cancer. Steinman's passion continued even when it was him in the bed, receiving dendritic-cell therapy of his own design. ■

**Michel C. Nussenzweig** is a Howard Hughes Medical Institute investigator at Rockefeller University. **Ira Mellman** is at Genentech and the University of California, San Francisco. Both trained with Steinman early in their scientific careers.  
e-mails: [nussen@mail.rockefeller.edu](mailto:nussen@mail.rockefeller.edu), [mellman.ira@gene.com](mailto:mellman.ira@gene.com)



D. H. JONES/SPL

**Figure 1 | Cold clouds.** Manney *et al.*<sup>2</sup> report that Arctic stratospheric temperatures were low enough for polar stratospheric clouds, such as the ones shown here over Swedish Lapland, to form through much of the winter of 2010–11. These clouds have a crucial role in polar ozone destruction.

## ATMOSPHERIC SCIENCE

# An Arctic ozone hole?

**The observation of unusually low ozone levels over the Arctic last winter provides reassuring evidence that our knowledge of stratospheric chemistry is robust. Whether such an episode will happen again is an open question. [SEE ARTICLE P.469](#)**

ROLANDO R. GARCIA

In late March of this year, the thickness of the ozone column in the atmosphere above the Arctic dropped to minimum values of 220–230 Dobson units (DU). Although these values are much higher than the lowest values seen routinely since the 1990s for the Antarctic spring (about 100 DU), they are nonetheless among the lowest ozone-column figures ever observed in the Arctic<sup>1</sup>. These values were recorded for just one week in late March, but minimum column amounts of less than 250 DU were observed for almost a month. Reporting on page 469 of this issue, Manney *et al.*<sup>2</sup> argue that the very low ozone observed over the Arctic in March constitutes the first ever ‘ozone hole’ in the Northern Hemisphere.

The authors support their argument by showing that the chemistry of the Arctic stratosphere in the spring of 2011 was remarkably similar to that commonly observed in Antarctica, which routinely leads to the development of the well-known Antarctic ozone hole. In particular, the abundances of nitric acid (HNO<sub>3</sub>) and hydrochloric acid (HCl) were exceptionally low in March, and close to values usually seen only in Antarctica (see Fig. 2 of the paper<sup>2</sup>, which shows observations of these

and other species, including ozone, made by the Microwave Limb Sounder Instrument on NASA’s Aura satellite). At the same time, the abundance of chlorine monoxide (ClO), which is an efficient catalyst of ozone destruction, rose throughout the Arctic winter and remained at record high levels during March — levels that are comparable to those seen in Antarctica in September, the equivalent point in the seasonal cycle.

The high levels of catalytic chlorine-containing species such as ClO led to rapid ozone loss over the Arctic from mid-February through into March. In mid-to-late March, the ozone abundance at altitudes near 20 kilometres dropped to about 1.5 parts per million by volume (p.p.m.v.). This is in the upper range of values observed in Antarctica in September, but well below typical values for the Arctic. By the end of March, the abundance of ozone was 1–1.5 p.p.m.v. at altitudes between 15 and 20 km; again, this resembles what happens during the Antarctic spring, and is much lower than normal for the Arctic.

Manney *et al.* show that the unusual evolution of atmospheric chemistry observed in the Arctic was made possible by the occurrence of persistently low stratospheric temperatures from December 2010 through to March 2011. There have been other exceptionally cold

Arctic winters in the recent past: 1995–96, 1999–2000 and 2004–05 were as cold as last winter from December through to February, and 1996–97 was just as cold during March and April. But only in 2010–11 were very cold conditions sustained continuously for a period of four months. Minimum temperatures 20 km above the Arctic remained below 195 K from mid-December 2010 through to the end of March 2011. This value is crucial, because it is the threshold for the formation of polar stratospheric clouds (PSCs; Fig. 1). PSC particles provide sites for heterogeneous (gas–solid and gas–liquid) reactions<sup>3</sup> that ‘activate’ chlorine by freeing it from the reservoir species HCl and chlorine nitrate (ClONO<sub>2</sub>). Indeed, widespread formation of PSCs was documented by Manney and colleagues<sup>2</sup>. This is consistent with their observations of high levels of ClO and low levels of HNO<sub>3</sub> and HCl, because heterogeneous reactions on PSCs deplete levels of HCl and ‘denitrify’ the stratosphere by removing HNO<sub>3</sub> from the gas phase.

Did ozone depletion in 2011 constitute a true Arctic ozone hole? As unusual as this year was, ozone loss did not approach the magnitude of depletion seen in Antarctica, whether measured locally or in terms of the total ozone column. In Antarctica, ozone virtually disappears at altitudes between 12 and 24 km, and in some

## BIOCHEMISTRY

# Suicide of a protein

**Plants and fungi follow a complex route to make the vitamin thiamine for carbohydrate metabolism. One of the pathway's protein participants turns out to be a surprising player, sacrificing its own activity in the process. SEE LETTER P.542**

PETER ROACH

years the depth of the ozone column is reduced to 100 DU or even less<sup>1</sup>. Such severe depletion was not observed in the Arctic. Judged by these criteria, there was no Arctic ozone hole in 2011 — only the most extreme episode of ozone loss seen in the Arctic so far. However, the evolution of HNO<sub>3</sub>, HCl and ClO species was strikingly Antarctic-like and different from what has been observed in other Arctic winters. On the basis of these considerations, Manney *et al.* conclude, reasonably enough, that there was an Arctic ozone hole in 2011.

Was the ozone loss seen in 2011 a truly exceptional event, or should we expect repeated episodes of extensive Arctic ozone depletion as a result of climate change? Although increasing concentrations of carbon dioxide in the atmosphere warm the troposphere, they cool the overlying stratosphere<sup>4</sup>, and observations<sup>5</sup> have shown that the global stratosphere has cooled significantly during the past 30 years. Furthermore, there is evidence to suggest that the coldest Arctic winters are becoming colder<sup>6</sup> — although there is no statistically significant trend because Arctic winter temperatures are highly variable. Comprehensive models of the climate system do not produce consistent predictions for the evolution of Arctic temperatures and ozone loss in the twenty-first century<sup>7</sup>; the refinement of such predictions will undoubtedly be an important topic of research in the immediate future. Another consideration is that the abundance of ozone-destroying chlorine- and bromine-containing species in the atmosphere is decreasing following the adoption of the Montreal Protocol, which bans substances that deplete the ozone layer<sup>8</sup>, such that the impact of cold winters on stratospheric ozone will lessen steadily in the future.

All things considered, a repeat of this year's episode is possible, but not likely. Nonetheless, the behaviour of Arctic ozone in 2011 is a nice demonstration that our basic understanding of stratospheric chemistry is robust, and that extensive ozone loss will occur wherever the right conditions are present. ■

**Rolando R. Garcia** is in the Atmospheric Chemistry Division, National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307-3000, USA.  
e-mail: rgarcia@ucar.edu

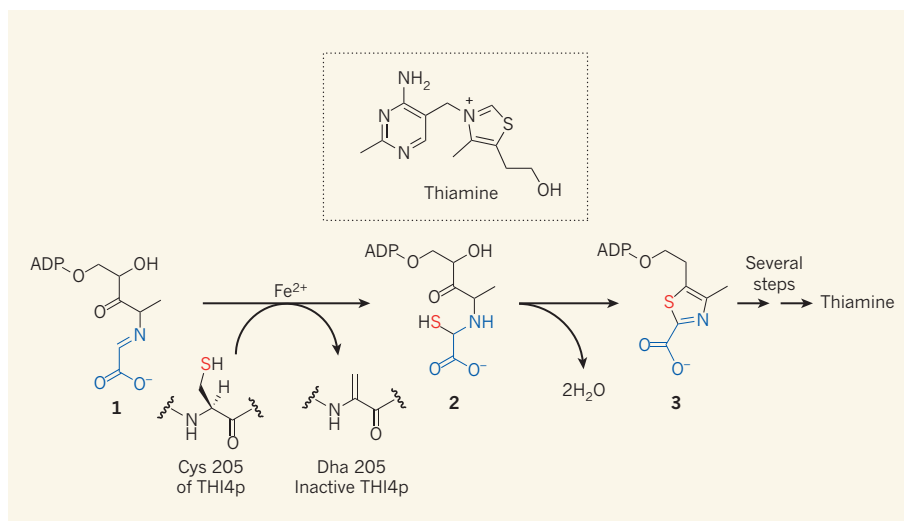
- Solomon, S., Portmann, R. W. & Thompson, D. W. J. *Proc. Natl Acad. Sci. USA* **104**, 445–449 (2007).
- Manney, G. L. *et al.* *Nature* **478**, 469–475 (2011).
- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755–758 (1986).
- Intergovernmental Panel on Climate Change *Contribution of Working Group I to the Fourth Assessment Report* (Cambridge Univ. Press, 2007).
- Randel, W. J. *et al.* *J. Geophys. Res.* <http://dx.doi.org/10.1029/2008JD010421> (2009).
- Tilmes, S., Müller, R., Engel, A., Rex, M. & Russell, J. M. *Geophys. Res. Lett.* <http://dx.doi.org/10.1029/2006GL026925> (2006).
- Charlton-Perez, A. J. *et al.* *Atmos. Chem. Phys.* **10**, 9473–9486 (2010).
- Jones, A. *et al.* *Atmos. Chem. Phys.* **11**, 5321–5333 (2011).

Thiamine is an essential cofactor that allows important enzymes to catalyse metabolic reactions. Humans depend on their diet to supply it as vitamin B<sub>1</sub>, but bacteria, plants and yeast can make their own. They do this by bringing together two precursor molecules: a sulphur-containing ring structure known as a thiazole and a nitrogenous pyrimidine. But the biochemical origin of the thiazole component of thiamine in plants and fungi has been mysterious — until now. On page 542 of this issue, Chatterjee and co-workers<sup>1</sup> describe a highly unusual reaction step in which a sulphur atom is transferred from a protein's amino-acid residue, to become part of the thiazole precursor. The donor protein is therefore functioning as a reagent, or 'suicide enzyme', rather than as a conventional catalytic enzyme.

The biosynthesis of the precursor thiazole starts with the metabolic cofactor molecule nicotinamide adenine dinucleotide (NAD); in the yeast *Saccharomyces cerevisiae*, the only protein required for the biosynthetic pathway is known as THI4p. The co-purification with

THI4p of compounds such as **3** (an adenylated carboxythiazole molecule; Fig. 1), which contains the same carbon skeleton as NAD, provided the first clues to NAD's function in this context<sup>2</sup>. Chatterjee *et al.*<sup>1</sup> found that one amino-acid residue — the cysteine at position 205 (Cys 205) of the THI4p sequence — loses a sulphur atom from its side chain during the biosynthesis, and so is converted to a dehydroalanine (Dha) residue. They concluded that the protein was probably providing the sulphur atom for thiazole biosynthesis.

Experimental evidence for a protein's role in a particular metabolic step is usually provided by reconstitution of the biosynthetic activity *in vitro* using the purified protein and a well-defined set of chemical ingredients. This was not easy to demonstrate for THI4p, because by the time it had been purified it had already lost the sulphur atom from Cys 205 — presumably as a result of undergoing reaction in the growing cells before purification. Chatterjee *et al.* got around the problem by culturing the cells on a growth medium containing the absolute minimum of nutrients: these cells yielded intact and unmodified THI4p protein. A paucity of reaction intermediates in



**Figure 1 | Incorporation of sulphur into a thiamine precursor.** The biosynthesis of the cofactor thiamine requires an adenylated carboxythiazole molecule (**3**) as a building block. Chatterjee *et al.*<sup>1</sup> report that, in yeast, the source of the sulphur atom in **3** is the cysteine amino-acid residue at position 205 of the protein THI4p. In the presence of iron(II) (Fe<sup>2+</sup>), Cys 205 reacts with precursor molecule **1** to afford intermediate compound **2**. The Cys 205 residue is converted into a dehydroalanine (Dha) residue in the process, thereby inactivating THI4p. Intermediate **2** then loses two molecules of water to produce **3**. This use of a protein as a co-substrate for a biosynthesis, rather than as a catalytic enzyme, is highly unusual. ADP is adenosyl diphosphate, a nucleotide group.

the medium after cell growth indicated that a component required for THI4p activity might be missing. However, addition of iron(II) ( $\text{Fe}^{2+}$ ) to the minimal-growth medium markedly increased the formation of compound 3 (Fig. 1). The reaction is extremely selective for  $\text{Fe}^{2+}$ , so the authors concluded that  $\text{Fe}^{2+}$  must be necessary for the sulphur-transfer reaction.

The thiazole-forming reaction could now be reconstituted *in vitro* from NAD using intact THI4p protein. The reaction showed conversion of NAD through to compound 1 (Fig. 1), and addition of  $\text{Fe}^{2+}$  completed the transformation to the final thiazole product, 3. Chatterjee *et al.* found that each molecule of THI4p accounted for the production of a single molecule of thiazole, observing a clear 1:1 correlation between the loss of sulphur from THI4p and formation of the thiazole product.

The deployment of a protein as such a metabolic reagent is extremely rare in biology. One

of the few well-characterized examples is the protein Ada, which repairs methylated lesions in DNA<sup>3</sup>. Reactivation of THI4p containing the Dha residue has so far proved impossible. Experiments with cultured cells show that thiamine and THI4p accumulate in a 1:1 ratio during growth, consistent with THI4p functioning as a 'one-shot' reagent and not being recycled by the cellular machinery. The reaction results in the build-up of inactivated (Dha 205) THI4p until it comprises about 1.5% of total cellular protein<sup>4</sup>.

These observations<sup>1</sup> raise interesting questions, such as why a protein should be used as a sulphur donor and why the inactive protein should not undergo rapid degradation and its amino acids be recycled. As Chatterjee *et al.* note, this accumulation of THI4p implies that the inactive protein has another cellular function. Indeed, inactivated THI4p has been implicated in DNA protection and other stress responses<sup>5–7</sup>. The biochemical mechanism linking the biosynthetic and stress-response

roles of THI4p is as yet unknown, but one possibility posed by Chatterjee *et al.* is that the protein helps to mop up any excess intracellular  $\text{Fe}^{2+}$ , thereby preventing chemical reactions that can produce damaging reactive oxygen species. ■

**Peter Roach** is at the School of Chemistry, University of Southampton, Southampton SO17 1BJ, UK.

e-mail: p.l.roach@soton.ac.uk

1. Chatterjee, A. *et al.* *Nature* **478**, 542–546 (2011).
2. Chatterjee, A., Jurgenson, C. T., Schroeder, F. C., Ealick, S. E. & Begley, T. P. *J. Am. Chem. Soc.* **129**, 2914–2922 (2007).
3. Sedgwick, B., Robins, P., Totty, N. & Lindahl, T. *J. Biol. Chem.* **263**, 4430–4433 (1988).
4. Faou, P. & Tropschug, M. *J. Mol. Biol.* **344**, 1147–1157 (2004).
5. Machado, C. R. *et al.* *J. Mol. Biol.* **273**, 114–121 (1997).
6. Ruiz-Roldán, C. *et al.* *Fungal Genet. Biol.* **45**, 6–16 (2008).
7. Medina-Silva, R. *et al.* *Res. Microbiol.* **157**, 275–281 (2006).

from Earth as Earth is from the Sun. Eris is the most distant Solar System object to be successfully observed by this technique, a notable achievement.

The authors find that Eris has a radius of  $1,163 \pm 6$  kilometres, which is smaller and has significantly less margin of error than the previous measurements of  $1,200 \pm 100$  km (ref. 3) and  $1,500 \pm 200$  km (ref. 4). Ironically, Sicardy and colleagues' results cannot definitively state whether Eris is larger than Pluto. The ambiguity is due to the fact that Pluto's atmosphere prevents accurate measurement of its surface location, rather than being due to any deficiency in the Eris observations. A more interesting result from the authors' study<sup>5</sup> is the possibility that Eris has a collapsed atmosphere (frozen to the surface), or a localized atmosphere under certain conditions.

Sicardy *et al.* demonstrate that Eris has no atmosphere at present and that a surprisingly large amount of light is reflected from its surface. This unusually bright surface is difficult to reconcile with the idea that objects in the outer Solar System become darkened by cosmic rays and micrometeorite impacts over time. The new observations could thus support a long-standing theory that, as a large, icy, Kuiper-belt object approaches the Sun during its orbit, a putative atmosphere could sublime and then condense out as the object moves farther away. Eris is currently far away from the Sun in its 557-year orbit, and although the results do not prove that an atmosphere ever existed, the bright surface could indicate relatively recent condensation.

The discovery of Eris and other Kuiper-belt objects allows Pluto, a seemingly unique object for so many years, to be placed in a broader perspective. Investigation of the bulk properties of a large number of Kuiper-belt objects

## PLANETARY SCIENCE

# Eris under scrutiny

**A stellar occultation by the dwarf planet Eris provides a new estimate of its size. It also reveals a surprisingly bright planetary surface, which could indicate the relatively recent condensation of a putative atmosphere. SEE LETTER P.493**

AMANDA GULBIS

**P**luto has long reigned as the smallest and most distant planet in our Solar System. In 1992, the discovery<sup>1</sup> of another body in roughly the same region proved that Pluto was not alone and launched a new frontier of study in planetary science — the Kuiper belt, the region at and beyond Neptune's orbit in which the two bodies reside. More than 1,000 Kuiper-belt objects have since been discovered. These objects consist mainly of ices and are typically located away from Earth at 30–50 times the distance between Earth and the Sun. One particular object, later named Eris, caught the attention of astronomers because initial size estimates based on proposed values for its surface reflectance showed that it was probably larger than Pluto<sup>2</sup>. Subsequent direct imaging suggested<sup>3</sup> that Eris was about 5% larger than Pluto, whereas detection of Eris's thermal radiation indicated<sup>4</sup> that the planet was substantially larger than Pluto. On page 493 of this issue, Sicardy *et al.*<sup>5</sup> use the powerful technique of stellar occultation to derive the most detailed information to date on Eris's size and other physical traits.

The discovery<sup>2</sup> of Eris in the Kuiper belt sparked the realization that more planets in the outer reaches of the Solar System were

probably awaiting detection. A great debate ensued as to whether Eris and other large, yet-to-be-detected Kuiper-belt objects — even Pluto itself — should be considered planets. This debate forced the astronomical community to rethink the definition of a planet. As a result, the definition was changed in 2006 and both Pluto and Eris were reclassified as dwarf planets. Rarely in the course of modern research has an astronomical discovery generated such widespread debate and emotional reaction among both scientists and the general public — it turns out that Eris was aptly named after the Greek goddess of strife and discord.

The stellar-occultation technique, in which a star is observed to pass behind a foreground object (in this case, Eris), has proven to be an effective method for discovering and characterizing features of Solar System bodies. Only a handful of large Kuiper-belt objects have been observed to occult stars, and each occultation has revealed something new and interesting, such as an unexpectedly bright surface on Kuiper-belt object 55636 (ref. 6), or waves in the upper atmosphere of Pluto<sup>7,8</sup>. These results have led to reconsideration of ideas about how objects in the outer Solar System are formed and evolve. Sicardy *et al.*<sup>5</sup> predicted and observed the occurrence of a stellar occultation when Eris was nearly 100 times as distant

has provided insight into the formation, evolution and dynamical histories of these bodies. We now expect that there are hundreds of objects that will eventually be classified as dwarf planets. Because Eris is the only body similar to Pluto for which detailed stellar-occultation data are available, Sicardy and colleagues' results<sup>5</sup> represent a major step forward in our knowledge about large Kuiper-belt objects.

Future attempts will certainly be made to

obtain additional stellar-occultation data for Eris as well as for other Kuiper-belt objects. Whether they are called planets or not, there is clearly still much to learn about these distant, icy bodies. ■

**Amanda Gulbis** is at the Southern African Large Telescope and South African Astronomical Observatory, 7935 Cape Town, South Africa.  
e-mail: amanda@salt.ac.za

1. Jewitt, D. & Luu, J. *Nature* **362**, 730–732 (1993).
2. Brown, M. E., Trujillo, C. A. & Rabinowitz, D. L. *Astrophys. J.* **635**, L97–L100 (2005).
3. Brown, M. E., Schaller, E. L., Roe, H. G., Rabinowitz, D. L. & Trujillo, C. A. *Astrophys. J.* **643**, L61–L63 (2006).
4. Bertoldi, F., Altenhoff, W., Weiss, A., Menten, K. M. & Thum, C. *Nature* **439**, 563–564 (2006).
5. Sicardy, B. *et al.* *Nature* **478**, 493–496 (2011).
6. Elliot, J. L. *et al.* *Nature* **465**, 897–900 (2010).
7. Person, M. J. *et al.* *Astron. J.* **136**, 1510–1518 (2008).
8. McCarthy, D. W. *et al.* *Astron. J.* **136**, 1519–1522 (2008).

## GENOMICS

# Plague's progress

**The Black Death was one of the most devastating pandemics in human history. The first complete genome sequence of the causative *Yersinia pestis* bacterium provides a fresh perspective on plague evolution. SEE LETTER P.506**

EDWARD C. HOLMES

**T**he Black Death was a pandemic of almost unprecedented scale. It is estimated that 30–50% of Europe's population perished from the plague between 1347 and 1351. As might be expected from such a remarkable level of mortality, this pestilence had a profound impact on medieval society. For example, it greatly influenced the depiction of death in art (Fig. 1), and acted as the driving force for the establishment of some of the first public-health measures, although these were generally futile. It has long been supposed that the Black Death, like all plague epidemics since, was caused by the bacterium *Yersinia pestis* in a transmission cycle involving fleas, rats and humans. But the evolutionary relationships between *Y. pestis* strains from different plague epidemics have been less clear. On page 506 of this issue, Bos *et al.*<sup>1</sup> describe the first complete genome sequence of *Y. pestis* from Black Death victims, and show that this pandemic was a pivotal event in plague evolution.

Obtaining bacterial DNA that is almost 700 years old presents a number of challenges, prominent among them the risk of inadvertent contamination by DNA from other sources<sup>1,2</sup>. Bos *et al.* extracted bacterial DNA from five teeth of plague victims taken from a burial pit at East Smithfield (ES) in London, which was established at the height of the pandemic in 1348–49. The key methodological advance in their work was the use of a molecular capture assay that assisted in the detection of *Y. pestis* DNA amid a background of host and environmental DNA. The stringent laboratory procedures used by the authors, the observed patterns of mutational damage in the DNA, and the finding that the ES strain is ancestral to all contemporary *Y. pestis* strains on



**Figure 1 | The last rites.** In this fourteenth-century picture, a victim of the Black Death is attacked by a devil, while a priest reads the last rites and God watches from above.

phylogenetic trees, strongly suggest that the obtained genome sequence is an authentic representative of the Black Death pathogen.

The ES strain is significant for several reasons. First, in combination with studies of partial genome sequences of *Y. pestis* from Black Death victims<sup>3</sup>, it demonstrates beyond doubt that *Y. pestis* was the true cause of the Black Death, despite claims to the contrary<sup>4</sup>. Second, the fact that the ES strain falls at the base of a phylogenetic tree that links contemporary *Y. pestis* genomes suggests that the Black Death was a crucial event in plague evolution that

generated all later lineages of *Y. pestis*, including those responsible for the 'Asian' (or modern) plague pandemic that has spread globally since the nineteenth century. Interestingly, the greatest genetic diversity of *Y. pestis* is observed<sup>5</sup> in China, suggesting that most plague epidemics originated in this region. Third, Bos *et al.*<sup>1</sup> suggest that the genetic similarity between the ES strain and contemporary strains of *Y. pestis* that are associated with less severe epidemics indicates that the high mortality of the Black Death was not simply a function of the bacterial strain involved. Given the small sample size of the authors' study, this assertion must be treated with caution, but the availability of a complete genome of *Y. pestis* from the Black Death should make the hypothesis experimentally testable.

Although the recovery and sequencing of the ES strain confirms the role of *Y. pestis* in the Black Death, it also raises questions about the cause of two earlier major disease pandemics previously assigned to *Y. pestis*: one that spread through parts of Africa, Asia and Europe in AD 541–542, during the reign of the Roman emperor Justinian and, more tentatively, the plague of Athens (430–426 BC), which was evocatively described in the writings of the Greek historian Thucydides. Bacterial DNA has purportedly been recovered from both of these epidemics, although in the case of the plague of Athens, the DNA was attributed<sup>6</sup> to the *Salmonella enterica* serovar Typhi bacterium that causes typhoid fever, rather than to *Y. pestis*. However, the Athenian *Salmonella* strain is not actually closely related to that responsible for typhoid<sup>7</sup>, suggesting that the ancient DNA had been contaminated by DNA from a modern, soil-living *Salmonella* species. The cause of the plague of Athens therefore remains a mystery.

The DNA supposedly from the Justinian plague is certainly that of *Y. pestis*, but the close similarity between the 'Justinian' DNA and that of modern *Y. pestis* variants<sup>8</sup> suggests that the former strain may not be authentic. Importantly, if all contemporary strains of *Y. pestis* are derived from the Black Death, as suggested by Bos and colleagues<sup>1</sup>, then both of the earlier epidemics were caused either by a strain of *Y. pestis* that has left no contemporary descendants, or by an entirely different organism. Ancient DNA may be central to resolving this question.

The analysis of ancient DNA is a powerful

THE BRITISH LIBRARY/HPH/TOFFOTO

tool, providing a unique perspective on the timescale and rates of microbial evolution, and of the host adaptation of microbial populations. Ancient DNA also provides a means to determine the cause of ancient epidemics, particularly when historical descriptions of disease symptoms yield ambiguous diagnoses. However, the spectre of contamination has loomed large, so that the impact of ancient-DNA analysis on our understanding of the evolution of infectious diseases has been relatively minor. Up until now, the most important study of historical pathogens has been the sequencing<sup>9</sup> of the virus responsible for the devastating influenza pandemic of 1918–19, which has provided insight into viral pathogenesis<sup>10</sup>. The

ES strain of *Y. pestis* is one of the oldest pathogens sequenced so far, and the results are some of the most credible in the field's history. Bos and colleagues' study<sup>1</sup> certainly sheds light on one of the most significant events in human history, but its greatest significance may be that it heralds a new era of research into the genomics of ancient bacterial diseases. ■

**Edward C. Holmes** is at the Center for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University, University Park, Pennsylvania 16802, USA, and the Fogarty International Center, Bethesda, Maryland.  
e-mail: echolmes@psu.edu

1. Bos, K. I. *et al. Nature* **478**, 506–510 (2011).
2. Gilbert, M. T. P. *et al. Microbiology* **150**, 341–354 (2004).
3. Haensch, S. *et al. PLoS Pathog.* **6**, e1001134 (2010).
4. Scott, S. & Duncan, C. J. *Biology of Plagues* (Cambridge Univ. Press, 2001).
5. Morelli, G. *et al. Nature Genet.* **42**, 1140–1143 (2010).
6. Papagrigorakis, M. J., Yapijakis, C., Synodinos, P. N. & Baziotopoulou-Valavani, E. *Int. J. Infect. Dis.* **10**, 206–214 (2006).
7. Shapiro, B., Rambaut, A. & Gilbert, M. T. P. *Int. J. Infect. Dis.* **10**, 334–335 (2006).
8. Drancourt, M. *et al. Emerg. Infect. Dis.* **10**, 1585–1592 (2004).
9. Taubenberger, J. K. *et al. Nature* **437**, 889–893 (2005).
10. Tumpey, T. M. *et al. Science* **310**, 77–80 (2005).

## ASTROPHYSICS

# Shedding light on the fabric of space–time

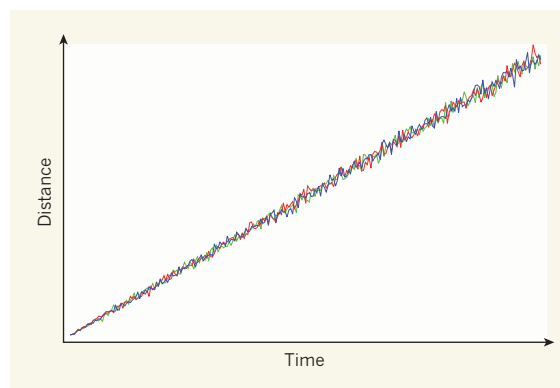
**The idea that space–time might be fundamentally fuzzy is much debated among theorists. A search for signatures of this effect on light from distant cosmic sources has come up empty-handed, but shows the potential of this approach.**

GIOVANNI AMELINO-CAMELIA

One of the most fascinating journeys in science concerns the evolution of the description of space and time. Examples of milestones on this journey are the realization that time is relative and that space-time bends in response to the presence of heavy particles and bodies<sup>1</sup>. Only one feature of the initial naive conceptualization of space and time by Isaac Newton has survived: in current theories, space-time is still viewed as a smooth entity. However, even this last Newtonian pillar is now being scrutinized. The main prediction of the theory of quantum mechanics, which underlies many of the advances of fundamental physics in the twentieth century, is that the results of a large class of measurements are affected by irreducible uncertainties. And all attempts to apply this successful theory to the description of space-time suggest that, as a result of some of these uncertainties, space-time should be fundamentally fuzzy. Writing in *Astronomy & Astrophysics*, Tamburini *et al.*<sup>2</sup> describe how they have searched for an imprint of space-time fuzziness on light from sources located at large distances from Earth.

In a fuzzy space-time, particles would travel much like a car on a bumpy road, with the notion of a smooth trajectory emerging only if observations cannot

resolve the effects of the bumps (Fig. 1). Unfortunately, the magnitude of the 'bumps' of space-time — the scale of space-time fuzziness — is horrifyingly small. This scale is expected to be of the order of the minuscule Planck length,  $L_p$ , which is about  $10^{-35}$  metres and is defined as  $L_p = (\hbar G/c^3)^{1/2}$ , where  $G$  is Newton's constant of gravity,  $\hbar$  is the reduced Planck constant of quantum mechanics and  $c$  is the speed of light.



**Figure 1 | Bumpy journeys.** The distance travelled by free particles in space increases with time, but the exact form of the increase depends on the nature of the fabric of space-time. In a smooth space-time, the increase is exactly linear, whereas in a 'bumpy' space-time the dependence is only roughly so. Shown here, qualitatively, are three particle journeys (blue, red and green) in a bumpy space-time. The effects of space-time bumps are invisibly small for short distances, but become visible for larger distances. Tamburini *et al.*<sup>2</sup> looked for evidence of a bumpy space-time on light from distant cosmic sources. (Figure courtesy of N. Loreti.)

Theoretical studies of space-time fuzziness have been an active area of research since at least the 1960s, when John Wheeler pointed out<sup>3</sup> the significance of this feature and coined the catchy term 'space-time foam' to describe it. But for a long time there was no experimental counterpart to these studies, because it seemed that 'bumps' with a magnitude of  $10^{-35}$  metres should forever be beyond the reach of the sensitivity of any experimental apparatus. Moreover, devising detailed candidate models of space-time fuzziness is hard, because essentially it involves combining the complexity of quantum mechanics with the intricacies of general relativity<sup>4</sup>.

As a result, phenomenological approaches to describing space-time foam started to gain momentum only quite recently. A first wave of studies<sup>5,6</sup> attempted to 'parameterize our ignorance' about space-time foam and to compensate for the smallness of the Planck length by studying the effects of fuzziness for particles travelling across distances of a few kilometres on Earth. But then, with a second wave of investigations<sup>7,8</sup>, it became clear that this strategy could be generalized to the case of observations of light from certain classes of distant astrophysical sources. And it is this astrophysics version of the approach to testing space-time fuzziness that Tamburini *et al.*<sup>2</sup> adopt in their study.

The authors' analysis exploits the fact that, as a result of space-time fuzziness, the wavefront of a light wave from a distant source is expected to develop tiny corrugations as it travels, such that some portions of the wavefront are advanced while others are retarded. For propagation over large cosmological distances from Earth, this wavefront corrugation can add up to give a macroscopic effect. Tamburini *et al.* searched for this effect in images of around 160 distant quasars (extremely luminous galactic nuclei powered by supermassive black holes), all at redshift greater than 4. Their analysis benefits from both the large number

of quasars and a calibration technique that is based on images of relatively nearby stars, for which the effects of space-time fuzziness are expected to be negligible.

The result of Tamburini and colleagues' analysis is negative: the authors found no evidence for space-time fuzziness. But this could be the start of an exciting season for space-time research. Their study convincingly establishes that, by adopting techniques based on wave-front corrugation, several parameterizations of space-time fuzziness occurring at the Planck length can already be ruled out and others could be tested in the near future. This was inconceivable in the days of Wheeler's proposal, and only a speculative hope even a few years ago<sup>5,6</sup>.

At this point, theories of space-time fuzziness might be lagging behind experiments. It is widely appreciated that the multi-parameter models adopted in studies such as that of Tamburini *et al.* are rather crude. These models should be good enough for preliminary estimates but are not sufficiently refined to

provide detailed guidance for experimentalists looking for a manifestation of space-time foam. In particular, there is no compelling mathematical description for the behaviour of light waves in the presence of fuzziness. So far, the descriptions rely on taking as a starting point Maxwell's wave equation and then introducing heuristically some plausible implications of fuzziness. This is partly why the authors of a study<sup>9</sup> similar to that of Tamburini *et al.*, while also concluding that this type of analysis can allow detection of fuzziness at the Planck scale, legitimately argue for slightly different estimates of the bounds presently established on some parameterizations of fuzziness.

The development of refined models of space-time foam does not look much easier now than it did in the 1960s. But having established the possibility of probing fuzziness at the Planck length experimentally, there will be plenty of motivation for facing the challenges that theory presents. If experimentalists are guided by refined descriptions

of space-time foam, they might find its signature in the not-so-distant future. But for now, the last pillar of the Newtonian description of space-time still stands. ■

**Giovanni Amelino-Camelia** is in the Department of Physics, Università di Roma La Sapienza, and the Sezione Roma 1, INFN, Piazzale Moro 2, Rome 00185, Italy. e-mail: giovanni.amelino-camelia@roma1.infn.it

1. Stachel, J. in *100 Years of Relativity* (ed. Ashtekar, A.) 3–36 (World Scientific, 2005).
2. Tamburini, F., Cuofano, C., Della Valle, M. & Gilmozzi, R. *Astron. Astrophys.* **533**, A71 (2011).
3. Wheeler, J. A. *Geometrodynamics* (Academic, 1962).
4. Carlip, S. *Rep. Prog. Phys.* **64**, 885–942 (2001).
5. Amelino-Camelia, G. *Nature* **398**, 216–218 (1999).
6. Ng, Y. J. & van Dam, H. *Found. Phys.* **30**, 795–805 (2000).
7. Lieu, R. & Hillman, L. W. *Astrophys. J.* **585**, L77–L80 (2003).
8. Ragazzoni, R., Turatto, M. & Gaessler, W. *Astrophys. J.* **587**, L1–L4 (2003).
9. Christiansen, W. A., Ng, Y. J., Floyd, D. J. E. & Perlman, E. S. *Phys. Rev. D* **83**, 084003 (2011).

## BIOPHYSICS

# More than a bystander

**The tendency of hydrophobic surfaces to aggregate in water is often invoked to explain how biomolecules recognize and bind to each other. Water seems to have a much more active role in these processes than had been thought.**

PHILIP BALL

When biomolecules interact, what do the surrounding water molecules do? One might think that their job would simply be to get out of the way, a crowd that must stand aside for the main actors. But there is now good reason to believe that water has a much more active role in the dialogue between the more celebrated constituents of the cell. When a protein binds its ligand, associates with another protein or folds into its functional form, the surrounding solvent acts as a versatile intermediary and facilitator. Three papers<sup>1–3</sup> uncover the subtlety and sophistication of that role and, in doing so, challenge some common perceptions of how biomolecular recognition operates.

One of the key concepts in the interactions of biomolecules is the hydrophobic effect, which loosely characterizes the tendency of hydrophobic particles and surfaces to aggregate in aqueous environments<sup>4,5</sup>. Proteins typically bury their hydrophobic amino-acid residues in their interior as they fold; and hydrophobic groups on ligands are generally juxtaposed to similar surfaces in an enzyme's binding site. Proteins themselves associate into larger aggregates — as functional assemblies, for example, or as fibrillar misfolded

structures in neurodegenerative diseases — by marrying up their hydrophobic surfaces. Yet there is still no consensus on how these hydrophobic interactions operate.

The traditional picture, now decades old, invokes an enhanced ordering of water molecules around the hydrophobe that preserves the hydrogen bonding of water molecules in the bulk<sup>6</sup>. In this view, the coming together of hydrophobes expels the intervening 'ordered' water into the bulk phase, an entropically favourable process. Although still routinely used, this picture receives no real support from experiments probing local hydration structures<sup>7</sup>.

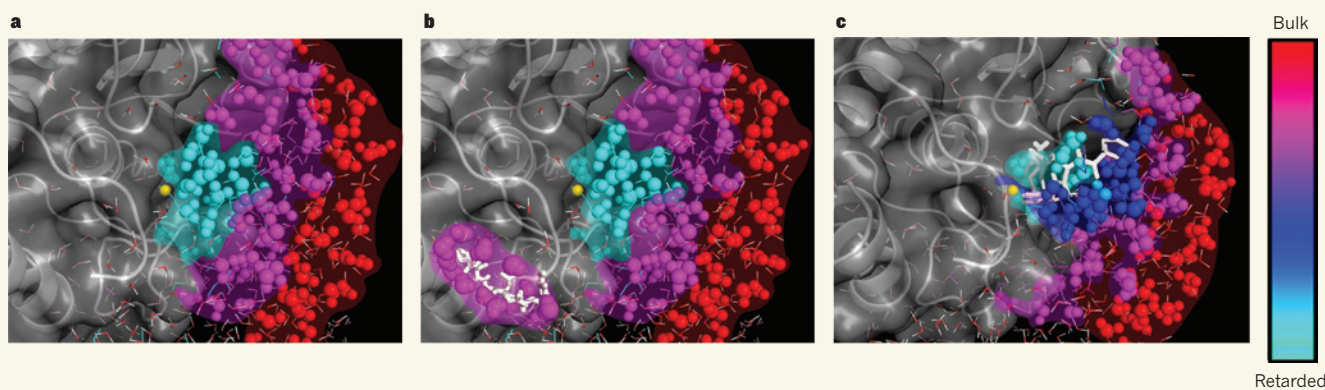
It has been argued<sup>8</sup> that there are, in fact, two distinct size regimes in which hydrophobic effects work. Small hydrophobic particles can be accommodated in water without affecting hydrogen-bonding networks, whereas larger ones unavoidably break hydrogen bonds, generating a 'soft' interface at the surface of the particle that is analogous to that between liquid and vapour. These extended surfaces then stick together because of a de-wetting transition: at a certain separation of particles, the water is collectively expelled from between them. The crossover between these two pictures is predicted to be at size scales of about 1 nanometre — precisely the scale that is relevant to the

association and folding of proteins. Evidence for de-wetting in the aggregation of protein subunits has, however, been conflicting<sup>9,10</sup>.

In *Proceedings of the National Academy of Sciences*, Li and Walker<sup>1</sup> now report that such a length-scale crossover exists, albeit one that decreases from around 11.4 Ångströms at 48 °C to 3.5 Å at 150 °C. They have used the tip of an atomic-force microscope to unravel a collapsed hydrophobic polymer in such a way that monomers buried in the collapsed mass become exposed to water one at a time, whereupon they become hydrated by water molecules. The authors found that, for polymers in which the monomers have large hydrophobic side-chains (about 1 nm long), there is a maximum value for the free energy of monomer solvation at which the entropy of hydration changes from being positive to negative — in other words, this is the crossover predicted previously<sup>8</sup>.

One objection to the de-wetting picture was that it was likely to be too difficult to nucleate a 'dry' region between hydrophobic surfaces — that is, to create a vapour cavity in the water. But a recently published model<sup>10</sup> of solvation suggests that part of the surface of melittin, a peptide that can apparently aggregate by de-wetting<sup>11</sup>, is sufficiently hydrophobic to permit spontaneous cavitation. Others have suggested<sup>12,13</sup> that de-wetting draws on the intrinsic fluctuations of water density at the water–hydrophobe interface, and another recent study<sup>14</sup> has shown that biomolecules may tune these fluctuations so that they sit close to a de-wetting transition. This tendency of biological systems to bring themselves in proximity to a phase transition, and thereby to enable sensitive and pronounced responses to small changes in the environment, is probably generic<sup>15</sup>.

Also in *Proceedings of the National Academy*



**Figure 1 | Water motion in enzyme–substrate binding.** Grossman *et al.*<sup>3</sup> report that the movements of water molecules associated with the binding site of a zinc metalloprotease enzyme change substantially during binding of a substrate to the enzyme. **a**, In this computer simulation of the free enzyme, the enzyme surface is grey and the zinc ion in the active site is yellow. Water molecules are shown as spheres, the colours of which indicate the mobility of the molecules, based on the timescale at which the hydrogen-bonded network of molecules rearranges: red indicates relatively free motion, as found in bulk water; cyan indicates strongly retarded motion; and colours in between

represent intermediate levels of motion. A steep gradient of water motion is observed from outside to inside the active site. **b**, In the early stage of binding, a substrate (white) is bound nonspecifically to the surface of the enzyme. The substrate has its own cohort of hydrating water molecules. **c**, Once the substrate is specifically bound to the zinc ion, the gradient of water motion around the site is far less steep than in **a**. The motion of the water molecules solvating the substrate is also slowed down compared with **b**. Grossman *et al.* propose that the change in water dynamics assists the binding of the substrate to the active site. (Taken from Fig. 5b of the paper<sup>3</sup>.)

of Sciences, Snyder *et al.*<sup>2</sup> show that the whole discourse of the hydrophobic effect, at least in ligand binding, has too long been dominated by the notion that there is a single explanation involving the expulsion of water from the binding cleft. As with any process, the free-energy change associated with ligand binding contains an entropic and an enthalpic contribution (enthalpy is a measure of the total energy of the system). During binding, water molecules constrained inside a cleft might be released, thereby boosting the free energy of binding by increasing entropy. But the enthalpic contribution to that change is by no means obvious, and could potentially counteract any entropic gain.

Snyder and colleagues' results show that it is probably unwise to make any generalizations about these thermodynamic contributions. They have characterized the binding between a rigid enzyme and a series of structurally related substrates. Some of these ligands contained groups that increase their hydrophobic contact area with the binding cavity, which has a hydrophobic and a hydrophilic side. The authors strikingly conclude that the alleged hydrophobic effect is rather insensitive to this contact area. Instead, it seems to arise primarily from structural changes in the network of water molecules between the ligand and the hydrophilic side of the cavity. Thus, although at a broad level the hydrophobic effect does involve differences in the structure of water close to solute surfaces relative to the structure of bulk water, the detailed balance of entropy and enthalpy is likely to vary on a case-by-case basis, and can be understood only by this kind of detailed analysis. Moreover, Snyder *et al.*<sup>2</sup> point out that "the shape of the water in the binding cavity may be as important as the shape of the cavity".

Although all this makes for a far more complicated picture of biomolecular binding than the classic geometrical 'lock and key' model, it is still predicated on a static or quasi-equilibrium picture. But that, too, is incomplete, according to Grossman and colleagues' report in *Nature Structural and Molecular Biology*<sup>3</sup>. They have used a combination of spectroscopy techniques, coupled to molecular-dynamics simulations, to follow changes in water and protein dynamics as a zinc metalloprotease enzyme binds its substrate. The results offer perhaps the most astonishing picture of how finely biomolecules manipulate their associated water molecules to perform their function.

The authors<sup>3</sup> find that, as enzyme–substrate binding develops, but before a full complex is formed, the movement of water near the protein is retarded (Fig. 1). Crudely put, it is as if the water 'thickens' towards a more glassy form, which in turn calms the fluctuations of the substrate so that it can become locked securely in place. It is not yet clear what causes this solvent slowdown as a precursor to binding; indeed, the whole question of cause and effect is complicated by the close coupling of protein and water motion and will be tricky to disentangle. In any event, molecular recognition here is much more than a case of complementarity between receptor and substrate — it also crucially involves the solvent. This suggests that changes in protein and solvent dynamics are not mere epiphenomena, but have a vital role in substrate binding and recognition: they are more cause than consequence.

As well as offering a fresh view of biomolecular shape and function, these findings<sup>1–3</sup> pose a daunting yet stirring challenge. Given that most drugs are ligands that bind to biological targets, will it be possible to make the fine-tuning of water structure and dynamics

an element of drug design? Indeed, can we hope to compete systematically with natural recognition processes at drug targets unless that mastery is attained? ■

**Philip Ball** is a writer based in London.

1. Li, I. T. S. & Walker, G. C. *Proc. Natl Acad. Sci. USA* **108**, 16527–16532 (2011).
2. Snyder, P. W. *et al. Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1114107108> (2011).
3. Grossman, M. *et al. Nature Struct. Mol. Biol.* **18**, 1102–1108 (2011).
4. Tanford, C. *The Hydrophobic Effect* 2nd edn (Wiley, 1980).
5. Southall, N. T., Dill, K. A. & Hayment, A. D. J. *J. Phys. Chem. B* **106**, 521–533 (2002).
6. Kauzmann, W. *Adv. Protein Chem.* **14**, 1–63 (1959).
7. Blokzijl, W. & Engberts, J. B. F. *N. Angew. Chem. Int. Edn* **32**, 1545–1579 (1993).
8. Lum, K., Chandler, D. & Weeks, J. D. *J. Phys. Chem. B* **103**, 4570–4577 (1999).
9. Zhou, R., Huang, X., Margulis, C. J. & Berne, B. J. *Science* **305**, 1605–1609 (2004).
10. Wang, J., Kudesia, S., Bratko, D. & Luzar, A. *Phys. Chem. Chem. Phys.* <http://dx.doi.org/10.1039/c1cp22082a> (2011).
11. Liu, P., Huang, X., Zhou, R. & Berne, B. J. *Nature* **437**, 159–162 (2005).
12. Huang, D. M. & Chandler, D. J. *J. Phys. Chem. B* **106**, 2047–2053 (2002).
13. Willard, A. P. & Chandler, D. J. *J. Phys. Chem. B* **112**, 6187–6192 (2008).
14. Patel, A. J. *et al.* preprint at <http://arxiv.org/abs/1109.4431> (2011).
15. Mora, T. & Bialek, W. *J. Stat. Phys.* **144**, 268–302 (2011).

#### CORRECTION

In the News & Views article 'Chemical biology: Many faces of a cancer-supporting protein' by John F. Darby and Paul Workman (*Nature* **478**, 334–335; 2011), the authors' declaration of competing financial interests was inadvertently omitted. Details can be found in the online article at [go.nature.com/x7hedv](http://go.nature.com/x7hedv)

# Unprecedented Arctic ozone loss in 2011

Gloria L. Manney<sup>1,2</sup>, Michelle L. Santee<sup>1</sup>, Markus Rex<sup>3</sup>, Nathaniel J. Livesey<sup>1</sup>, Michael C. Pitts<sup>4</sup>, Pepijn Veefkind<sup>5,6</sup>, Eric R. Nash<sup>7</sup>, Ingo Wohltmann<sup>3</sup>, Ralph Lehmann<sup>3</sup>, Lucien Froidevaux<sup>1</sup>, Lamont R. Poole<sup>8</sup>, Mark R. Schoeberl<sup>9</sup>, David P. Haffner<sup>7</sup>, Jonathan Davies<sup>10</sup>, Valery Dorokhov<sup>11</sup>, Hartwig Gernandt<sup>3</sup>, Bryan Johnson<sup>12</sup>, Rigel Kivi<sup>13</sup>, Esko Kyrö<sup>13</sup>, Niels Larsen<sup>14</sup>, Pieter F. Levelt<sup>5,6,15</sup>, Alexander Makshtas<sup>16</sup>, C. Thomas McElroy<sup>10</sup>, Hideaki Nakajima<sup>17</sup>, Maria Concepción Parrondo<sup>18</sup>, David W. Tarasick<sup>10</sup>, Peter von der Gathen<sup>3</sup>, Kaley A. Walker<sup>19</sup> & Nikita S. Zinoviev<sup>16</sup>

**Chemical ozone destruction occurs over both polar regions in local winter–spring. In the Antarctic, essentially complete removal of lower-stratospheric ozone currently results in an ozone hole every year, whereas in the Arctic, ozone loss is highly variable and has until now been much more limited. Here we demonstrate that chemical ozone destruction over the Arctic in early 2011 was—for the first time in the observational record—comparable to that in the Antarctic ozone hole. Unusually long-lasting cold conditions in the Arctic lower stratosphere led to persistent enhancement in ozone-destroying forms of chlorine and to unprecedented ozone loss, which exceeded 80 per cent over 18–20 kilometres altitude. Our results show that Arctic ozone holes are possible even with temperatures much milder than those in the Antarctic. We cannot at present predict when such severe Arctic ozone depletion may be matched or exceeded.**

Since the emergence of the Antarctic ‘ozone hole’ in the 1980s<sup>1</sup> and elucidation of the chemical mechanisms<sup>2–5</sup> and meteorological conditions<sup>6</sup> involved in its formation, the likelihood of extreme ozone depletion over the Arctic has been debated. Similar processes are at work in the polar lower stratosphere in both hemispheres, but differences in the evolution of the winter polar vortex and associated polar temperatures have in the past led to vastly disparate degrees of spring-time ozone destruction in the Arctic and Antarctic. We show that chemical ozone loss in spring 2011 far exceeded any previously observed over the Arctic. For the first time, sufficient loss occurred to reasonably be described as an Arctic ozone hole.

## Arctic polar processing in 2010–11

In the winter polar lower stratosphere, low temperatures induce condensation of water vapour and nitric acid (HNO<sub>3</sub>) into polar stratospheric clouds (PSCs). PSCs and other cold aerosols provide surfaces for heterogeneous conversion of chlorine from longer-lived reservoir species, such as chlorine nitrate (ClONO<sub>2</sub>) and hydrogen chloride (HCl), into reactive (ozone-destroying) forms, with chlorine monoxide (ClO) predominant in daylight<sup>5,7</sup>.

In the Antarctic, enhanced ClO is usually present for 4–5 months (through to the end of September)<sup>8–11</sup>, leading to destruction of most of the ozone in the polar vortex between ~14 and 20 km altitude<sup>7</sup>. Although ClO enhancement comparable to that in the Antarctic occurs at some times and altitudes in most Arctic winters<sup>9</sup>, it rarely persists for more than 2–3 months, even in the coldest years<sup>10</sup>. Thus chemical ozone loss in the Arctic has until now been limited, with largest previous losses observed in 2005, 2000 and 1996<sup>7,12–14</sup>.

The 2010–11 Arctic winter–spring was characterized by an anomalously strong stratospheric polar vortex and an atypically long continuously cold period. In February–March 2011, the barrier to

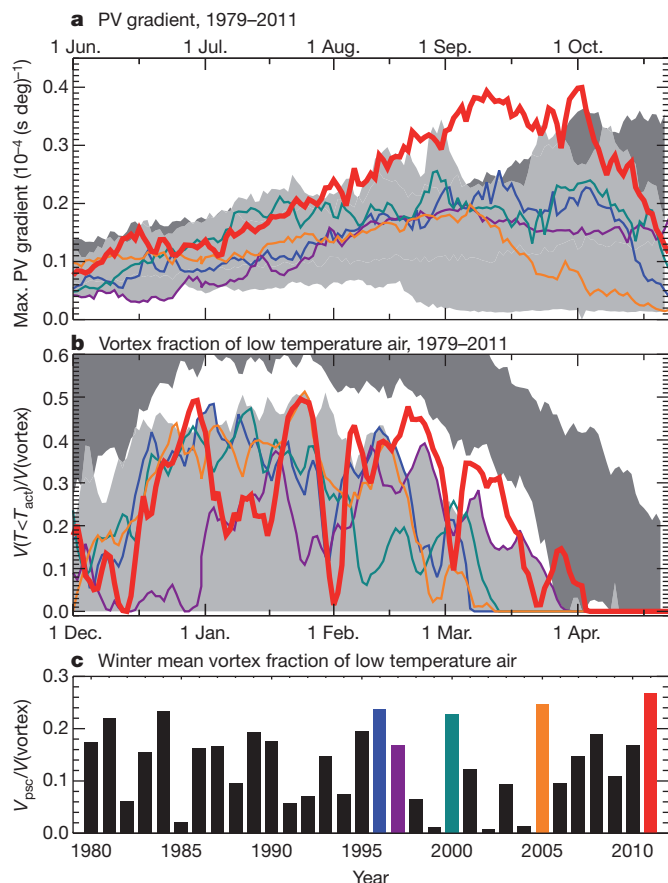
transport at the Arctic vortex edge was the strongest in either hemisphere in the last ~30 years (Fig. 1a, Supplementary Discussion).

The persistence of a strong, cold vortex from December through to the end of March was unprecedented. In the previous years with most ozone loss, temperatures ( $T$ ) rose above the threshold associated with chlorine activation ( $T_{\text{act}}$ , near 196 K, roughly the threshold for the potential existence of PSCs) by early March (Fig. 1b, Supplementary Figs 1, 2). Only in 2011 and 1997 have Arctic temperatures below  $T_{\text{act}}$  persisted through to the end of March, sporadically approaching a vortex volume fraction similar in size to that in some Antarctic winters (Fig. 1b). In 1996–97, however, the cold volume remained very limited until mid-January and was smaller than that in 2011 at most times during late January through to the end of March (Fig. 1b, Supplementary Figs 1, 2).

Daily minimum temperatures in the 2010–11 Arctic winter were not unusually low, but the persistently cold region was remarkably deep (Supplementary Figs 1, 2). Temperatures were below  $T_{\text{act}}$  for more than 100 days over an altitude range of ~15–23 km, compared to a similarly prolonged cold period over only ~20–23 km altitude in 1997; below ~19 km altitude,  $T < T_{\text{act}}$  continued for ~30 days longer in 2011 than in 1997 (Supplementary Fig. 1b). In 2005, the previous year with largest Arctic ozone loss<sup>7</sup>,  $T < T_{\text{act}}$  occurred for more than 100 days over ~17–23 km altitude, but all before early March.

The winter mean volume of air in which PSCs may form (that is, with  $T < T_{\text{act}}$ ),  $V_{\text{psc}}$ , is closely correlated with the potential for ozone loss<sup>7,15–17</sup>. In 2011,  $V_{\text{psc}}$  (as a fraction of the vortex volume) was the largest on record (Fig. 1c). Both large  $V_{\text{psc}}$  and cold lingering well into spring are important in producing severe chemical loss<sup>7,15,16</sup>, and 2010–11 was the only Arctic winter during which both conditions have been met. Much lower fractional  $V_{\text{psc}}$  in 1997 than in 1996, 2000, 2005 or 2011 (Fig. 1c) is consistent with less ozone loss that year<sup>16,17</sup>.

<sup>1</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA. <sup>2</sup>New Mexico Institute of Mining and Technology, Socorro, New Mexico 87801, USA. <sup>3</sup>Alfred Wegener Institute for Polar and Marine Research, D-14473 Potsdam, Germany. <sup>4</sup>NASA Langley Research Center, Hampton, Virginia 23681, USA. <sup>5</sup>Royal Netherlands Meteorological Institute, 3730 AE De Bilt, The Netherlands. <sup>6</sup>Delft University of Technology, 2600 GA Delft, The Netherlands. <sup>7</sup>Science Systems and Applications, Inc., Lanham, Maryland 20706, USA. <sup>8</sup>Science Systems and Applications, Inc., Hampton, Virginia 23666, USA. <sup>9</sup>Science and Technology Corporation, Lanham, Maryland 20706, USA. <sup>10</sup>Environment Canada, Toronto, Ontario, Canada M3H 5T4. <sup>11</sup>Central Aerological Observatory, Dolgoprudny 141700, Russia. <sup>12</sup>NOAA Earth System Research Laboratory, Boulder, Colorado 80305, USA. <sup>13</sup>Arctic Research Center, Finnish Meteorological Institute, 99600 Sodankylä, Finland. <sup>14</sup>Danish Climate Center, Danish Meteorological Institute, DK-2100 Copenhagen, Denmark. <sup>15</sup>Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands. <sup>16</sup>Arctic and Antarctic Research Institute, St Petersburg 199397, Russia. <sup>17</sup>National Institute for Environmental Studies, Tsukuba-city, 305-8506, Japan. <sup>18</sup>National Institute for Aerospace Technology, 28850 Torrejón De Ardoz, Spain. <sup>19</sup>University of Toronto, Toronto, Ontario, Canada M5S 1A7.



**Figure 1 | Meteorology of the Arctic lower stratosphere.** **a**, Vortex strength (as indicated by maximum potential vorticity<sup>49</sup> (PV) gradients) at 460 K potential temperature ( $\sim 18$  km altitude,  $\sim 65$  hPa level). **b**, Fraction of vortex volume at potential temperatures between 390 and 550 K with a temperature less than the chlorine activation threshold ( $T_{act}$ ). Light (dark) grey shading shows range of Arctic (Antarctic) values for 1979–2010. Antarctic dates are shifted by six months (top axis in **a**) to show the equivalent season. **c**, Winter mean  $V_{psc}$  during the past 32 years, expressed as a fraction of vortex volume. Red, orange, green, purple and blue lines/bars show the 2010–11, 2004–05, 1999–2000, 1996–97 and 1995–96 Arctic winters, respectively.

Factors playing secondary parts in governing interannual variability in ozone destruction, including vortex strength, structure and position relative to the cold region, also favour large loss in 2011 (Supplementary Figs 2, 3, Supplementary Discussion). However, despite the fraction of the vortex with  $T < T_{act}$  and mid-March temperatures sporadically approaching those seen in the Antarctic (Fig. 1b, Supplementary Fig. 1a), even in 2011 temperatures were much higher, and the cold regions much smaller, than those in most Antarctic winters.

Satellite trace-gas and PSC measurements highlight the stark contrast between polar processing in 2010–11 and that in typical Arctic winters, and the parallels with Antarctic conditions (Figs 2, 3). In 2011, PSCs or aerosols were abundant until mid-March (Fig. 3a; consistent with a deep region with  $T < T_{act}$ , Fig. 3b), much later than usual in the Arctic<sup>18–20</sup>, with vortex-average amounts at some altitudes similar to those in the Antarctic and dramatically larger than the near-zero values at that time in most Arctic winters. Furthermore, PSCs in 2011 spanned an altitude range comparable to that in the Antarctic, an uncommon occurrence in the Arctic<sup>18–20</sup>. Particles in long-lasting PSCs can grow large enough to sediment, resulting in denitrification, permanent removal of  $\text{HNO}_3$  from the stratosphere<sup>7,12</sup>. By late March 2011 no PSCs remained (Fig. 3a), yet  $\text{HNO}_3$  mixing ratios were much lower than observed in any previous Arctic winter (Fig. 2a). The continuing depression in  $\text{HNO}_3$  after PSCs had evaporated indicates

denitrification. Albeit less severe than in typical Antarctic winters (Fig. 2b, c, 3c), the extent and degree of denitrification in 2011 were unmatched in the Arctic, approaching the range of Antarctic conditions for the first time.

Decreasing HCl and increasing ClO signify chlorine activation (Fig. 2d–i). Some ClO enhancement has occurred in all recent Arctic winters, but has never been as prolonged and extensive as that in 2011. In late February, high ClO pervaded the sunlit portion of the vortex. The 2011 values vastly exceed the range previously observed in the Arctic from late February through to the end of March. They also briefly lie outside the Antarctic seasonal envelope, primarily because the higher solar zenith angles of the Antarctic measurements used here lead to  $\sim 30\%$  lower ClO under fully activated conditions. In late February, HCl values (unaffected by solar zenith angle issues) fall along the lower boundary of the Antarctic envelope, confirming the picture seen in ClO. The vertical extent of chlorine activation was also comparable to that in the Antarctic (Fig. 3d, e).

In previous cold Arctic winters, chlorine was deactivated (converted from ozone-destroying forms into less reactive reservoir species) by mid-March<sup>11</sup>; even in 1997, ClO started to decline by late February (Fig. 2g). In 2011, by contrast, ClO began decreasing rapidly only about a week earlier than is typical in the Antarctic. ClO data in late February 1997 indicate that not only were maximum values lower than those in early March 2011, but also the vertical range of enhancement was shallower, with weaker activation at low altitudes than in 2011 (Fig. 3e), consistent with the higher altitudes and decreasing extent (Figs 1b, 3b, Supplementary Fig. 2) of  $T < T_{act}$ .

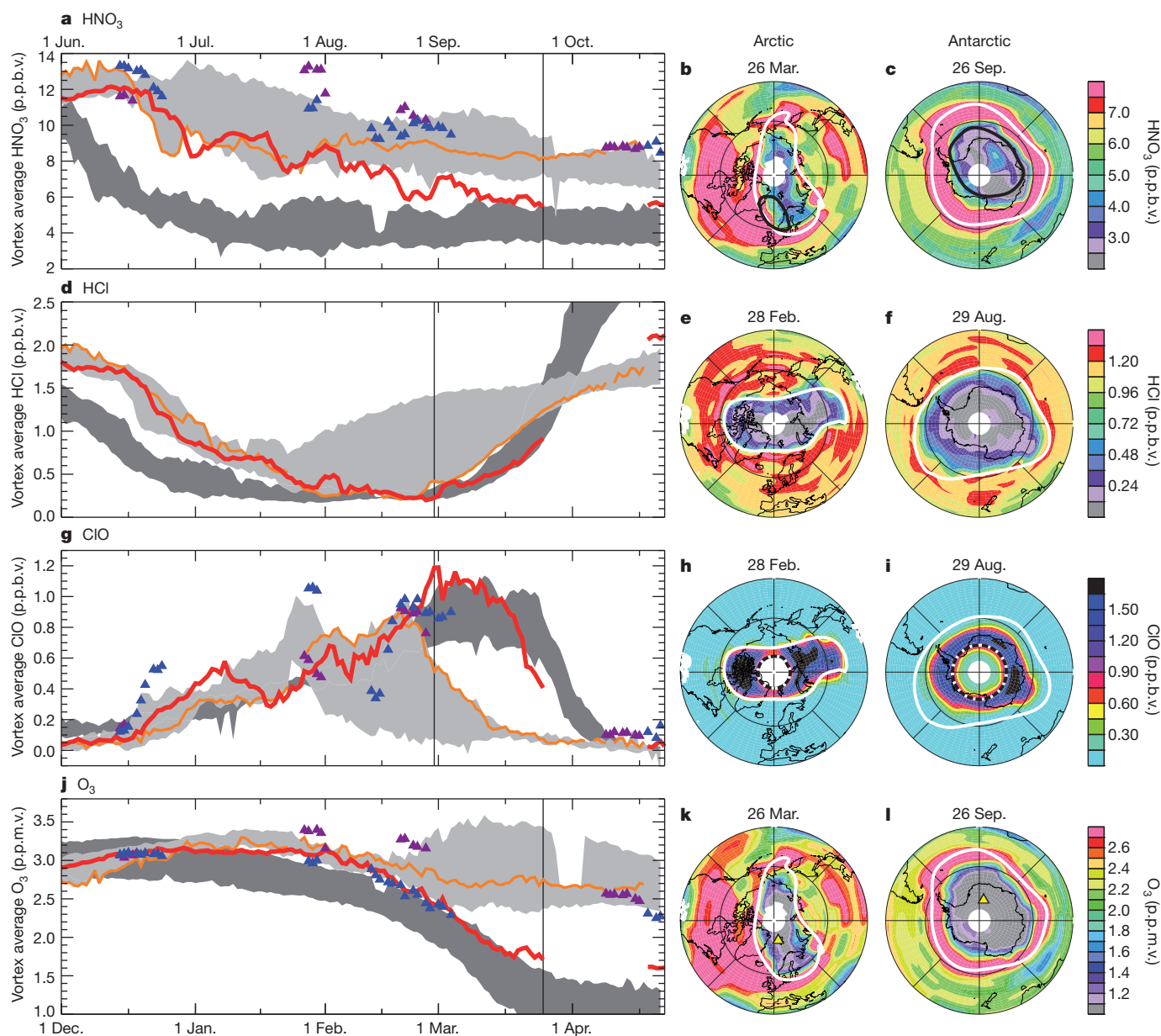
When chlorine is deactivated, whether it is converted first into HCl or  $\text{ClONO}_2$  depends sensitively upon  $\text{HNO}_3$  and ozone abundances. In the Arctic, chlorine is normally deactivated through initial reformation of  $\text{ClONO}_2$ . In the severely denitrified and ozone-depleted Antarctic vortex, production of  $\text{ClONO}_2$  is suppressed and that of HCl highly favoured<sup>11,12,21</sup>. In March 2011, the recovery of HCl followed a much more Antarctic-like pathway than has been observed in any other Arctic winter.

The largest Arctic chemical ozone loss was previously observed in 2005, followed closely by 2000 and 1996<sup>7,12–14</sup>. Although low temperatures persisted until the end of March 1997, the ozone loss in that year was far less. No previous year rivals 2011, when the evolution of Arctic ozone more closely followed that typical of the Antarctic (Fig. 2j). Ozone profiles in late March 2011 resemble typical Antarctic late-winter profiles much more strongly than they do the average Arctic one (Fig. 3f). Because mixing in April 2011 (for example, lamination events larger than that shown in Fig. 3f) entrained ozone-rich air into the vortex, the slight decrease in vortex-averaged ozone at a potential temperature of 485 K from 26 March to 20 April (from  $\sim 1.8$  to  $\sim 1.6$  p.p.m.v., Fig. 2j) indicates continuing chemical loss during this interval.

### Estimates of chemical ozone loss

Chemical loss is difficult to quantify in the Arctic, where transport from above replenishes ozone in the lower stratospheric vortex, obscuring the signature of chlorine-catalysed destruction<sup>12,22,23</sup>. The evolution of the long-lived trace gas nitrous oxide ( $\text{N}_2\text{O}$ ) reflects steady downward transport throughout the 2010–11 winter–spring, indicating that subsidence partially masked chemical loss. Horizontal transport can also confound the signature of chemical loss, bringing air into the vortex that has either higher<sup>24</sup> or lower<sup>14</sup> concentrations of ozone, depending on the altitude and latitude from which it originates.

Representative results from two types of chemical loss calculations<sup>24–28</sup> based on balloon-borne and satellite observations are shown in Fig. 4. The differences (up to  $\sim 0.4$  p.p.m.v. at the end of March 2011) in estimates derived from the various methods and data sets imply some uncertainty in the chemical loss determination. Year-to-year differences in the amount of ozone loss are very similar when obtained from any method/data set combination, however, indicating a high degree of precision in the relative amount of calculated loss



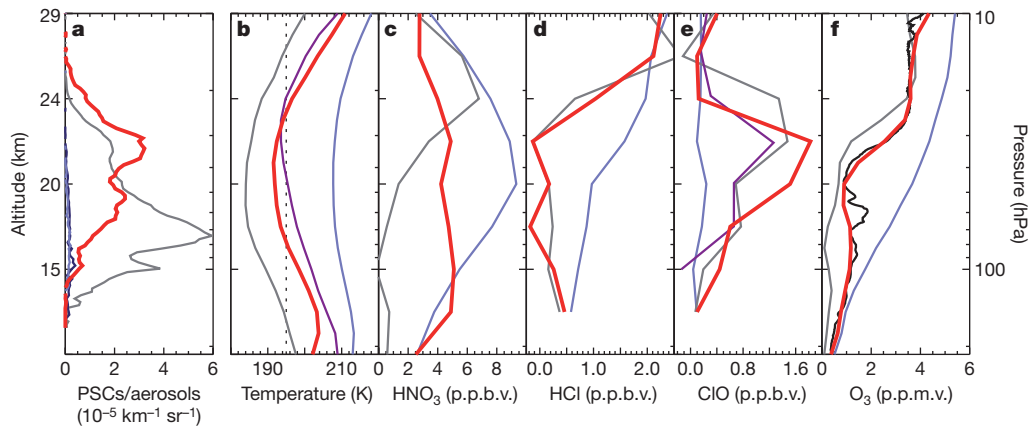
**Figure 2 | Chemical composition in the lower stratosphere.** a–l, Maps (right) and vortex-averaged time series (left) at 485 K potential temperature ( $\sim 20$  km,  $\sim 50$  hPa) for four different gases:  $\text{HNO}_3$  (a, b, c),  $\text{HCl}$  (d, e, f),  $\text{ClO}$  (g, h, i) and  $\text{O}_3$  (ozone; j, k, l); mixing ratios from Aura MLS are shown. Averaging for the time series is done within the white contour shown on the maps. Blue (purple) triangles on time series, 1995–96 (1996–97) values from UARS MLS. Line colours/shading as in Fig. 1, but shading is for Aura MLS measurements from

2005–10. Antarctic dates are shifted by six months (top axis on time series) to show the equivalent season. Vertical lines show dates of maps in 2011 (2010) in the Arctic (Antarctic). Black overlays on  $\text{HNO}_3$  maps,  $T_{\text{act}}$  ( $\sim 196$  K at this level);  $\text{HNO}_3$  may be sequestered in PSCs at lower temperatures. Dotted black/white contour on  $\text{ClO}$  maps,  $92^\circ$  SZA, poleward of which measurements were taken in darkness. Yellow/black triangles on ozone maps, locations of the profiles in Fig. 3.

between different years. Chemical destruction was severe between  $\sim 16$  and  $22$  km altitude, with the largest loss exceeding  $2.5$  p.p.m.v. by 26 March 2011 (Fig. 4a). By 31 March 2011, chemical loss was nearly double that in 2005 from  $\sim 18$  to above  $22$  km, and similar to that in 2005 at lower altitudes (Fig. 4b, c). From  $\sim 18$  to  $20$  km, more than 80% of the ozone present in January had been chemically destroyed by late March. Chemical removal in 1996 and 2000 started at a rate similar to that in 2011 (Fig. 4c), but ceased by late March; maximum losses in 2000 approached those in 2011, but extended over a much smaller vertical range (Fig. 4b). Loss in 1996, 2000 and 2005 considerably exceeded that in 1997, with greater destruction at lower altitudes in those years contributing more to total column loss<sup>7,12,13</sup>. Chemical loss in 2011 was two to three times larger than that in 1997, and about twice that in 1996 and 2005 above  $\sim 16$  km; from  $\sim 15$  to  $23$  km it was comparable to that in the Antarctic ozone hole in 1985<sup>29</sup>.

Single ozone-sonde station measurements in early April 2011 suggest continuing ozone loss (Fig. 4c).

Although the meteorology during March–April was similar in 1997 and 2011, ozone loss was much more pronounced in 2011. Photochemical box model simulations (Supplementary Fig. 4, Supplementary Discussion) elucidate how early winter conditions set the stage for record springtime ozone destruction in 2011. Chlorine activation brought on by enduring cold from December through to the end of February led to  $\sim 0.7$ – $0.8$  p.p.m.v. lower ozone at the beginning of March 2011 (Figs 2j, 4c). The early onset of continuous cold also facilitated formation of PSC particles large enough to sediment, resulting in  $\sim 4$  p.p.b.v. less  $\text{HNO}_3$  by March in 2011 than in 1997 (Fig. 2a). The degree of denitrification has a profound impact on the severity of springtime Arctic ozone loss<sup>30</sup>. By delaying chlorine deactivation, lower  $\text{HNO}_3$  by 1 March was responsible for  $\sim 0.6$  p.p.m.v. more ozone

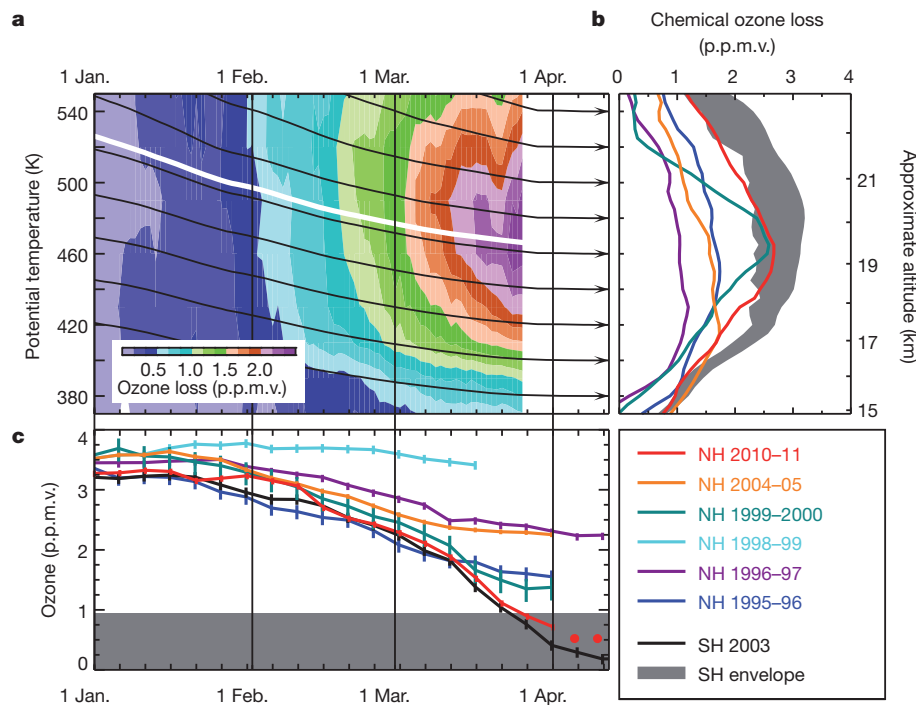


**Figure 3 | Vertical composition information.** **a**, Red, PSCs/aerosol amounts averaged in the vortex over a week centred around 25 February 2011; dark blue, the average for the same week in 2007–10; grey, the average over the equivalent period (centred on 28 August) for the Antarctic in 2006–10; lavender, the Arctic average for a week centred around 26 March 2011. (In late winter–spring, maximum PSC altitudes are generally higher in the Arctic because early winter PSC activity redistributes  $\text{HNO}_3$  and water vapour to lower altitudes in the Antarctic<sup>18</sup>). **b–f**, Daily average profiles of MERRA temperatures (**b**) and MLS  $\text{HNO}_3$  (**c**),  $\text{HCl}$  (**d**),  $\text{ClO}$  (**e**) and ozone (**f**). Red lines, data from a  $4^\circ \times 15^\circ$  latitude  $\times$  longitude box around  $79^\circ \text{N}$ ,  $12^\circ \text{E}$ ; in **c**, **f**, taken on 26 March; in

**b**, **d**, **e**, on 6 March 2011. Lavender, 7-day average for 2005–10 (1980–2010 for **b**) centred on the same location and days. Grey, profiles in a similar box in the Antarctic ( $79^\circ \text{S}$ ,  $12^\circ \text{E}$ ) on 26 September for **c**, **f**, and on 8 September 2010 for **b**, **d**, **e**. Dotted black line in **b**, approximate  $T_{\text{act}}$  (195 K), see text. Purple line in **b**, 7-day average around 6 March 1997, centred on same location. Purple line in **e**, a midday  $\text{ClO}$  profile from UARS MLS on 26 February 1997 averaged in an  $8^\circ \times 30^\circ$  box centred at the same Arctic location. A high-resolution ozone-sonde profile at Ny Ålesund on 26 March 2011 (black in **f**) agrees well with MLS; lamination, a signature of mixing with ozone-rich extra-vortex air, is apparent as a local maximum near 60 hPa.

loss after that date in 2011 than in 1997 (Supplementary Fig. 4, Supplementary Discussion). The effects of denitrification and early-winter loss together account for the disparity in ozone depletion in these two winters ( $\sim 1.5$  p.p.m.v. more loss at 460 K in 2011 than in 1997, Fig. 4c, Supplementary Fig. 4). Loss as severe as that in 2011 thus

requires  $T < T_{\text{act}}$  with consequent chlorine activation and ozone destruction, early in winter (as in 1996, 2000 and 2005, but not in 1997), a cold period and region before March sufficient to allow widespread denitrification, and the persistence of a cold polar vortex into April (as in 1997, but not in 1996, 2000 or 2005).

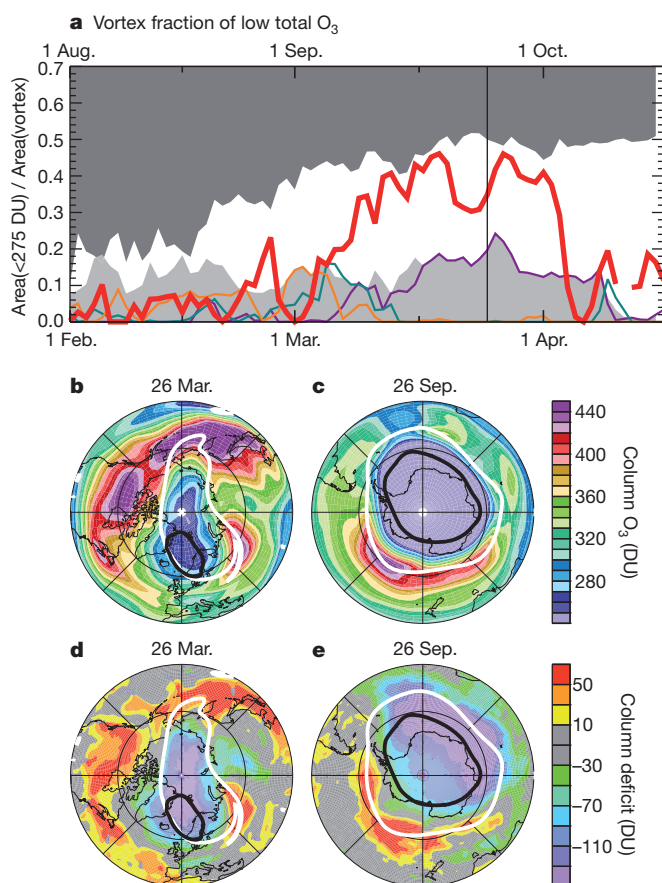


**Figure 4 | Chemical ozone loss estimates.** **a**, Chemical loss as a function of time and potential temperature from passive subtraction of MLS and ATLAS passively-transported ozone (initialized with December MLS data). **b**, Chemical loss from ozone sondes in unmixed vortex air as a function of 'spring equivalent potential temperature'<sup>48</sup> (black contours in **a**). Shading, Antarctic range defined by 1985 (the first year with profile measurements inside the ozone hole<sup>29</sup>) and 2003 (a recent year with a severe ozone hole). The 2003 Antarctic curve is shifted by six months minus 10 days because ozone sondes that year predominantly sampled the outermost vortex, where ozone

loss begins earliest. **c**, Ozone at a spring equivalent potential temperature of 465 K (white contour in **a**), near the level of maximum chemical loss. Shading, the region below the minimum reached in the 1985 Antarctic ozone hole. In April 2011 most soundings sampled the disturbed vortex edge; only two were made in air uninfluenced by mixing (red dots). Error bars,  $1\sigma$  uncertainties based on the scatter of individual ozone-sonde measurements. Line colours as in Fig. 1; 1998–99 (a winter with no ozone loss) is shown in cyan. NH, Northern Hemisphere; SH, Southern Hemisphere.

## Column ozone

Total column ozone is a predominant factor determining exposure of Earth's surface to ultraviolet radiation<sup>7,12</sup>. In the context of previous Arctic winters, 2011 was truly remarkable: the fraction of the Arctic vortex in March with total ozone less than 275 Dobson units (DU) is typically near zero, but reached nearly 45% in 2011 (Fig. 5a). Because of the dynamically-driven correlation between total ozone and lower-stratospheric temperature<sup>23,31–34</sup> (Supplementary Discussion), the abiding cold in 1997 and 2011 would have led to lower March total ozone than in other Arctic winters even without chemical loss; dynamical conditions in March–April 1997 particularly favoured low total ozone<sup>33</sup> (Supplementary Discussion). In March 2011, however, the area of low total ozone covered more than twice as much of the vortex as in 1997, and the daily vortex 'ozone deficit' (Supplementary Fig. 5a) was 30–50 DU larger, consistent with the greater chemical loss (Fig. 4). Maximum 2011 vortex fractions of low ozone approached those in early Antarctic ozone holes (Fig. 5a). The close correspondence between the vortex and both low total ozone and the large Arctic total ozone deficit (Fig. 5b, d) implies that low total ozone in March 2011 resulted primarily from chemical loss<sup>31,32</sup> (Supplementary Discussion). The ozone deficit in the Antarctic (Fig. 5e) shows a maximum over 0–90° W, and a minimum over 90–200° E, reflecting a vortex position in 2010 different to that in the reference state (which is less robust than that for the Arctic). Differences in morphology deep in the vortex are, however, minimal. The 2011 Arctic ozone deficit was at least comparable to that in the 2010 Antarctic vortex core at an equivalent time.



**Figure 5 | Total column ozone.** **a**, Time series of the fraction of 460 K vortex area with total ozone below 275 Dobson units (DU) in February–April in the Arctic (bottom axis), and in August–October in the Antarctic (top axis). Line colours/shading as in Fig. 1. 2005–2011 values are from OMI; earlier values are from TOMS (Total Ozone Mapping Spectrometer) instruments<sup>50</sup>. Maps show OMI total ozone (**b**, **c**) and ozone deficit (**d**, **e**) in the Arctic (Antarctic) on 26 March 2011 (26 September 2010). Overlays as in Fig. 2 but at 460 K.

## An echo of the Antarctic

In the absence of chemical ozone loss, downward transport during winter results in a springtime maximum in total ozone; because this transport is stronger in the Arctic, background ozone levels there are ~100 DU higher than those in the Antarctic<sup>7,23</sup>. Therefore Arctic spring total ozone could, even after chemical destruction comparable to that in an Antarctic ozone hole (commonly defined by values less than 220 DU; refs 7, 12), exhibit only a weak maximum in total ozone rather than a well-defined minimum. Examination of the long-term ozone-sonde record in the Arctic shows that abundances near 250 DU or less are well below typical autumn values, thus appearing as a 'hole' in total ozone. Dynamical processes can result in transient regions of very low total ozone (Supplementary Discussion, Supplementary Figs 5, 6) and/or local minima in lower-stratospheric ozone profiles (for example, via ozone-poor extra-vortex air transported into the polar vortex<sup>14,24</sup>). For an interhemispheric comparison of chemical loss, it is thus important to verify that observed Arctic ozone decreases were primarily related to chemical, rather than dynamical, processes.

Figure 4 shows that the precipitous decline in Arctic ozone in February–March 2011 resulted from chemical loss of similar magnitude to that in the Antarctic in the mid-1980s. Observed ozone between ~15 and 20 km altitude decreased to values matching the minima in early Antarctic ozone holes and those reached at the corresponding time in some recent Antarctic winters (Figs 2j–l; 3f). In late March–early April, most ozone-sonde profiles in the vortex had mixing ratios less than 1 p.p.m.v., with values ~0.7 p.p.m.v. over an approximately 2-km altitude region, and some dipping to 0.5 p.p.m.v. (Supplementary Fig. 7). Minimum total ozone in spring 2011 was continuously below 250 DU for ~27 days (Supplementary Fig. 5b), with a maximal area below that level of ~2 × 10<sup>6</sup> km<sup>2</sup> (roughly five times the area of Germany or California). Values dropped to ~220–230 DU for about a week in late March 2011.

In these respects, chemical ozone destruction in the 2011 Arctic polar vortex attained, for the first time, a level clearly identifiable as an Arctic ozone hole. On the other hand, although the magnitude of chemical depletion was comparable to that in the Antarctic, total ozone values remained higher and, because the areal extent of the Arctic vortex was much smaller (~60% the size of a typical Antarctic vortex), the low-ozone region was more confined.

The Arctic winter stratosphere exhibits striking interannual variability. The past decade has included the four most dynamically active (hence among the warmest) Arctic winters in the past 32 years (ref. 35) and now the two coldest winters with largest ozone loss<sup>7,12–14</sup>, extending the previously noted trend of the coldest winters becoming colder<sup>13,16</sup>. Had implementation of the Montreal Protocol not curbed the increase in stratospheric halogen loading, formation of an Arctic ozone hole would have already become common even in moderately cold winters<sup>36</sup>. Even with the lower anthropogenic halogen levels actually reached, the potential for Antarctic-like ozone loss in the Arctic in the event of a persistently cold winter–spring such as that in 2010–11 has been recognized for decades<sup>5,22</sup>. Despite temperatures that were generally far higher than those in Antarctic winter, Arctic chemical ozone destruction in 2011 rivalled that in some Antarctic ozone holes. The development of an Arctic ozone hole under conditions only slightly more extreme than those in some previous Arctic winters raises the possibility of yet more severe depletion as lower-stratospheric temperatures decrease. More acute Arctic ozone destruction could exacerbate biological risks from increased ultraviolet radiation exposure, especially if the vortex shifted over densely populated mid-latitudes, as it did in April 2011.

Our present understanding of what drives variability in the Arctic winter stratosphere is incomplete. Stratospheric temperatures and vortex evolution depend on the atmosphere's radiative properties and propagation of wave activity<sup>37,38</sup>, which are being modified by increasing greenhouse gas concentrations. Day-to-day tropospheric disturbances can lead to stratospheric warming or cooling, depending

on their geographical location and the stratospheric vortex structure, which controls their upward propagation<sup>39,40</sup>. Current climate models do not fully capture either the observed short-timescale patterns of Arctic variability or the full extent of the observed longer-term cooling trend in cold stratospheric winters; nor do they agree on future circulation changes that affect trends in transport<sup>41,42</sup>. Our ability to predict when conditions similar to, or more extreme than, those in 2011 may be realized is thus very limited. Improving our predictive capabilities for Arctic ozone loss, especially while anthropogenic halogen levels remain high, is one of the greatest challenges in polar ozone research. Comprehensive stratospheric data sets, such as those used here, are critical to meeting that challenge.

## METHODS SUMMARY

MERRA (Modern Era Retrospective-analysis for Research and Applications<sup>43</sup>) fields are used for temperature and vortex analysis and for vortex averaging of composition measurements. The CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) on the CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) satellite<sup>44</sup> provides PSC/aerosol information.

Trace gas profiles are from the Microwave Limb Sounder (MLS) on NASA's Aura satellite. Only daytime CIO measurements are used. Northern (southern) high latitudes are sampled near midday (in late afternoon), thus the average solar zenith angle (SZA) of MLS Antarctic measurements is  $\sim 7^\circ$  higher than that in the Arctic. Reactive chlorine partitioning shifts away from CIO at higher SZAs<sup>7,12</sup>, leading to  $\sim 30\%$  lower CIO measured in the Antarctic than in the Arctic under fully activated conditions. An instrument anomaly disrupted MLS measurements from 27 March to 20 April 2011. UARS (Upper Atmosphere Research Satellite) MLS measurements, used for 1995–1996 and 1996–1997 analyses, are sparse because of the UARS yaw cycle and other measurement gaps<sup>26</sup>.

Total column ozone is measured by the Dutch-Finnish Ozone Monitoring Instrument (OMI)<sup>46</sup> on Aura. Total ozone 'deficit' is the difference between daily values and a reference that is minimally affected by chemical loss.

Measurements from MLS and the Match network of balloon-borne ozone soundings (ozone sondes)<sup>47</sup> are used to estimate chemical ozone loss in two ways. The difference between calculated 'passive' (influenced only by transport) ozone and observed ozone is computed, with passive ozone obtained using MLS nitrous oxide<sup>14</sup>, a 'reverse trajectory' model<sup>25,26</sup>, and the ATLAS (Alfred Wegener Institute Lagrangian Chemistry/Transport System) model<sup>27</sup>. Vortex ozone is also examined on the surfaces on which it subsides<sup>12,14,28,48</sup>, with descent rates from modelled radiative heating/cooling rates averaged over the polar vortex<sup>48</sup>.

Photochemical box model runs were performed using the chemical model from ATLAS<sup>27</sup> to test the sensitivity of ozone loss to initial ozone amounts and denitrification.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 3 May; accepted 7 September 2011.

Published online 2 October 2011.

- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. Large losses of total ozone in Antarctica reveal seasonal  $\text{ClO}_x/\text{NO}_x$  interaction. *Nature* **315**, 207–210 (1985).
- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. On the depletion of Antarctic ozone. *Nature* **321**, 755–758 (1986).
- Molina, L. T. & Molina, M. J. Production of  $\text{Cl}_2\text{O}_2$  from the self-reaction of the CIO radical. *J. Phys. Chem.* **91**, 433–436 (1987).
- Anderson, J. G., Brune, W. H. & Proffitt, M. H. Ozone destruction by chlorine radicals within the Antarctic vortex: the spatial and temporal evolution of CIO- $\text{O}_3$  anticorrelation based on *in situ* ER-2 data. *J. Geophys. Res.* **94**, 11465–11479 (1989).
- Solomon, S. Stratospheric ozone depletion: a review of concepts and history. *Rev. Geophys.* **37**, 275–316 (1999).
- Schoeberl, M. R. & Hartmann, D. L. The dynamics of the stratospheric polar vortex and its relation to springtime ozone depletions. *Science* **251**, 46–52 (1991).
- World Meteorological Organization. *Scientific Assessment of Ozone Depletion: 2010* (Report 52, Global Ozone Research and Monitoring Project, 2011).
- Solomon, P. M. *et al.* High concentrations of chlorine monoxide at low altitudes in the Antarctic spring stratosphere: secular variation. *Nature* **328**, 411–413 (1987).
- Waters, J. W. *et al.* Stratospheric CIO and ozone from the Microwave Limb Sounder on the Upper Atmosphere Research Satellite. *Nature* **362**, 597–602 (1993).
- Santee, M. L., Manney, G. L., Waters, J. W. & Livesey, N. J. Variations and climatology of CIO in the polar lower stratosphere from UARS Microwave Limb Sounder measurements. *J. Geophys. Res.* **108**, 4454, <http://dx.doi.org/10.1029/2002JD003335> (2003).

- Santee, M. L. *et al.* A study of stratospheric chlorine partitioning based on new satellite measurements and modeling. *J. Geophys. Res.* **113**, D12307, <http://dx.doi.org/10.1029/2007JD009057> (2008).
- World Meteorological Organization. *Scientific Assessment of Ozone Depletion: 2006* (Report 50, Global Ozone Research and Monitoring Project, 2007).
- Rex, M. *et al.* Arctic winter 2005: implications for stratospheric ozone loss and climate change. *Geophys. Res. Lett.* **33**, L23808, <http://dx.doi.org/10.1029/2006GL026731> (2006).
- Manney, G. L. *et al.* EOS MLS observations of ozone loss in the 2004–2005 Arctic winter. *Geophys. Res. Lett.* **33**, L04802, <http://dx.doi.org/10.1029/2005GL024494> (2006).
- Harris, N. R. P., Lehmann, R., Rex, M. & von der Gathen, P. A closer look at Arctic ozone loss and polar stratospheric clouds. *Atmos. Chem. Phys.* **10**, 8499–8510 (2010).
- Rex, M. *et al.* Arctic ozone loss and climate change. *Geophys. Res. Lett.* **31**, L04116, <http://dx.doi.org/10.1029/2003GL018844> (2004).
- Tilmes, S., Müller, R., Engel, A., Rex, M. & Russell, J. M. III. Chemical ozone loss in the Arctic and Antarctic stratosphere between 1992 and 2005. *Geophys. Res. Lett.* **33**, L20812, <http://dx.doi.org/10.1029/2006GL026925> (2006).
- Poole, L. R. & Pitts, M. C. Polar stratospheric cloud climatology based on Stratospheric Aerosol Measurement II observations from 1978 to 1989. *J. Geophys. Res.* **99**, 13083–13089 (1994).
- Fromm, M. D. *et al.* An analysis of Polar Ozone and Aerosol Measurement (POAM) II Arctic stratospheric cloud observations, 1993–1996. *J. Geophys. Res.* **104**, 24341–24357 (1999).
- Pitts, M. C., Poole, L. R. & Thomason, L. W. CALIPSO polar stratospheric cloud observations: second-generation detection algorithm and composition discrimination. *Atmos. Chem. Phys.* **9**, 7577–7589 (2009).
- Douglas, A. R. *et al.* Interhemispheric differences in springtime production of HCl and CIONO<sub>2</sub> in the polar vortices. *J. Geophys. Res.* **100**, 13967–13978 (1995).
- Manney, G. L. *et al.* Chemical depletion of ozone in the Arctic lower stratosphere during winter 1992–93. *Nature* **370**, 429–434 (1994).
- Tegtmeier, S., Rex, M., Wohltmann, I. & Krüger, K. Relative importance of dynamical and chemical contributions to Arctic wintertime ozone. *Geophys. Res. Lett.* **35**, L17801, <http://dx.doi.org/10.1029/2008GL034250> (2008).
- Rex, M. *et al.* *In situ* measurements of stratospheric ozone depletion rates in the Arctic winter of 1991/1992: a Lagrangian approach. *J. Geophys. Res.* **103**, 5843–5853 (1998).
- Manney, G. L. *et al.* Lagrangian transport calculations using UARS data. Part II: ozone. *J. Atmos. Sci.* **52**, 3069–3081 (1995).
- Manney, G. L. *et al.* Variability of ozone loss during Arctic winter (1991–2000) estimated from UARS Microwave Limb Sounder measurements. *J. Geophys. Res.* **108**, 4149, <http://dx.doi.org/10.1029/2002JD002634> (2003).
- Wohltmann, I., Lehmann, R. & Rex, M. The Lagrangian chemistry and transport model ATLAS: simulation and validation of stratospheric chemistry and ozone loss in the winter 1999/2000. *Geosci. Model Dev.* **3**, 585–601 (2010).
- von der Gathen, P. *et al.* Observational evidence for chemical ozone depletion over the Arctic in winter 1991–92. *Nature* **375**, 131–134 (1995).
- Gernandt, H. The vertical ozone distribution above the GDR-research base, Antarctica in 1985. *Geophys. Res. Lett.* **14**, 84–86 (1987).
- Rex, M. *et al.* Prolonged stratospheric ozone loss in the 1995–96 Arctic winter. *Nature* **389**, 835–838 (1997).
- Manney, G. L., Froidevaux, L., Santee, M. L., Zurek, R. W. & Waters, J. W. MLS observations of Arctic ozone loss in 1996–97. *Geophys. Res. Lett.* **24**, 2697–2700 (1997).
- Manney, G. L., Santee, M. L., Froidevaux, L., Waters, J. W. & Zurek, R. W. Polar vortex conditions during the 1995–96 Arctic winter: meteorology and MLS ozone. *Geophys. Res. Lett.* **23**, 3203–3206 (1996).
- Petzoldt, K. The role of dynamics in total ozone deviations from their long-term mean over the Northern Hemisphere. *Ann. Geophys.* **17**, 231–241 (1999).
- Hood, L. L., Soukharev, B. E., Fromm, M. & McCormack, J. P. Origin of extreme ozone minima at middle to high northern latitudes. *J. Geophys. Res.* **106**, 20925–20940 (2001).
- Manney, G. L. *et al.* Aura Microwave Limb Sounder observations of dynamics and transport during the record-breaking 2009 Arctic stratospheric major warming. *Geophys. Res. Lett.* **36**, L12815, <http://dx.doi.org/10.1029/2009GL038586> (2009).
- Newman, P. A. *et al.* What would have happened to the ozone layer if chlorofluorocarbons (CFCs) had not been regulated? *Atmos. Chem. Phys.* **9**, 2113–2128 (2009).
- Newman, P. A., Nash, E. R. & Rosenfield, J. E. What controls the temperatures of the Arctic stratosphere during the spring? *J. Geophys. Res.* **106**, 19999–20010 (2001).
- Polvani, L. M. & Saravanan, R. The three-dimensional structure of breaking Rossby waves in the polar wintertime stratosphere. *J. Atmos. Sci.* **57**, 3663–3685 (2000).
- Orsolini, Y. J., Karpechko, A. Y. & Nikulin, G. Variability of the Northern Hemisphere polar stratospheric cloud potential: the role of North Pacific disturbances. *Q. J. R. Meteorol. Soc.* **135**, 1020–1029 (2009).
- Woollings, T., Charlton-Perez, A., Ineson, S., Marshall, G. & Masato, G. Associations between stratospheric variability and tropospheric blocking. *J. Geophys. Res.* **115**, D06108, <http://dx.doi.org/10.1029/2009JD012742> (2010).
- Butchart, N. *et al.* Multimodel climate and variability of the stratosphere. *J. Geophys. Res.* **116**, D05102, <http://dx.doi.org/10.1029/2010JD014995> (2011).
- Charlton-Perez, A. *et al.* The potential to narrow uncertainty in projections of stratospheric ozone over the 21st century. *Atmos. Chem. Phys.* **10**, 9473–9486 (2010).
- Reinecker, M. M. *et al.* MERRA — NASA's modern-era retrospective analysis for research and applications. *J. Clim.* **24**, 3624–3648 <http://dx.doi.org/10.1175/JCLI-D-11-00015.1> (2011).

44. Hunt, W. H. *et al.* CALIPSO lidar description and performance assessment. *J. Atmos. Ocean. Technol.* **26**, 1214–1228 (2009).
45. Waters, J. W. *et al.* The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite. *IEEE Trans. Geosci. Rem. Sens.* **44**, 1075–1092 (2006).
46. Levelt, P. F. *et al.* The Ozone Monitoring Instrument. *IEEE Trans. Geosci. Rem. Sens.* **44**, 1093–1101 (2006).
47. Rex, M. *et al.* Chemical ozone loss in the Arctic winter 1994/95 as determined by the Match technique. *J. Atmos. Chem.* **32**, 35–59 (1999).
48. Rex, M. *et al.* Chemical depletion of Arctic ozone in winter 1999/2000. *J. Geophys. Res.* **107**, 8276, <http://dx.doi.org/10.1029/2001JD000533> (2002).
49. Hoskins, B. J., McIntyre, M. E. & Robertson, A. W. On the use and significance of isentropic potential-vorticity maps. *Q. J. R. Meteorol. Soc.* **111**, 877–946 (1985).
50. McPeters, R. D. *et al.* *Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide* (NASA Technical Publication 1998-206895, 1998).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the MLS (especially A. Lambert, D. Miller, W. Read, M. Schwartz, P. Stek, J. Waters), OMI (especially P. K. Bhartia, G. Jaross, G. Labow), CALIPSO and Match science teams, as well as A. Douglass, J. Joiner and the Aura project, for their support. We also thank W. Daffer and R. Fuller for programming assistance at JPL; the many observers whose work went into obtaining the ozone-sonde measurements; the ozone scientists who participated in the discussion of the 2011 Arctic ozone loss and appropriate definition of an Arctic ozone hole (including, but not limited to, N. Harris, G. Bodeker, G. Braathen, M. Kurylo, R. Salawitch); and especially P. Newman and K. Minschwaner for discussions and comments. Meteorological analyses were provided by NASA's Global Modeling and Assimilation Office (GMAO) and by the European Centre for Medium-Range Weather Forecasts. We thank S. Pawson of GMAO for advice on usage of the MERRA reanalysis. Ozone-sonde measurements at Alert, Eureka, Resolute Bay, Churchill and Goose Bay were funded by Environment Canada. Additional ozone sondes were flown at Eureka as part of the

Canadian Arctic Atmospheric Chemistry Experiment (ACE) Validation Campaign and were funded by the Canadian Space Agency. Academy of Finland provided partial funding for performing and processing ozone-sonde measurements in Jokioinen and Sodankylä. Ozone soundings and work at AWI were partially funded by the EC DG Research through the RECONCILE project. Work at the Jet Propulsion Laboratory, California Institute of Technology, and at Science Systems and Applications Inc., was done under contract with NASA.

**Author Contributions** G.L.M. and M.L.S. led analysis of MLS data; M.R. led analysis of ozone-sonde data; G.L.M. led the meteorological data analysis. M.R., G.L.M., N.J.L. and I.W. did chemical ozone loss calculations. R.L. and M.R. performed and analysed chemical box model calculations. M.C.P. and L.R.P. provided CALIPSO/CALIOp data analyses; E.R.N. and P.V. provided TOMS and OMI data analyses. L.F., M.L.S., G.L.M. and N.J.L. provided expertise on MLS data usage; D.P.H., P.V. and P.F.L. provided expertise on OMI data usage. J.D., V.D., H.G., B.J., R.K., E.K., N.L., A.M., C.T.M., H.N., M.C.P., D.W.T., P.v.d.G., K.A.W. and N.S.Z. were responsible for performing and processing ozone-sonde measurements. All authors contributed comments on the manuscript. G.L.M., M.L.S. and M.R. jointly compiled and synthesized the results. G.L.M. and M.L.S. wrote the paper.

**Author Information** CALIOp data are publicly available at [http://eosweb.larc.nasa.gov/PRODOCS/calipso/table\\_calipso.html](http://eosweb.larc.nasa.gov/PRODOCS/calipso/table_calipso.html), MLS data at <http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/MLS>, OMI data at [http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omto3\\_v003.shtml](http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omto3_v003.shtml), and GEOS-5 MERRA analyses through <http://disc.sci.gsfc.nasa.gov/mdisc/data-holdings/merra/>. The balloon-borne Antarctic ozone-sonde data recorded in 1985 and the following years are publicly available at <http://dx.doi.org/10.1594/PANGAEA.547983>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to G.L.M. (Gloria.L.Manney@jpl.nasa.gov) or M.L.S. (Michelle.L.Santee@jpl.nasa.gov).

## METHODS

**Data sets.** Modern Era Retrospective-analysis for Research and Applications (MERRA)<sup>43</sup> fields, from the Goddard Earth Observing System Version 5.2.0 (GEOS-5) data assimilation system, are used for the temperature and vortex analysis. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite<sup>44</sup> provides PSC/aerosol information. CALIOP measurements began in April 2006. Trace gas profile measurements are from the Microwave Limb Sounder (MLS)<sup>45</sup> on NASA's Aura satellite, and the predecessor MLS instrument<sup>26</sup> on the Upper Atmosphere Research Satellite (UARS). Total column ozone data are from the Dutch-Finnish Ozone Monitoring Instrument (OMI)<sup>46</sup> on board Aura. The historical total ozone record comprises data from Nimbus-7 and Earth Probe Total Ozone Mapping Spectrometer (TOMS)<sup>50</sup>. Aura MLS and OMI measurements are available from August 2004 through to the present. UARS MLS measurements were obtained from September 1992 through to early 2000, with increasingly sparse sampling in the later years<sup>26</sup>. TOMS data are available beginning in 1979, but no TOMS instrument was taking measurements during the 1995–96 Arctic winter.

Measurements from the Match network of balloon-borne ozone soundings (ozone sondes)<sup>47</sup> are used in some of the chemical ozone loss estimates.

**Temperature and vortex analysis.** Potential vorticity<sup>49</sup> (PV) is used to define the vortex, with a contour of 'scaled' PV of  $1.4 \times 10^{-4} \text{ s}^{-1}$  (in vorticity units) demarking the vortex edge<sup>51,52</sup>. Vortex strength is diagnosed as the maximum daily gradient in PV as a function of equivalent latitude (the latitude that would enclose the same area between it and the pole as a given PV contour)<sup>51–53</sup>. Scaled PV multiplied by  $10^4$  is used in the calculation, resulting in units for its gradient of  $10^{-4} \text{ (s degrees equivalent latitude)}^{-1}$ .

The temperature threshold for chlorine activation,  $T_{\text{act}}$ , is estimated using the formula for nitric acid trihydrate formation<sup>54</sup>, which depends on pressure,  $\text{HNO}_3$  and  $\text{H}_2\text{O}$ . Climatological  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  profiles are used, derived from UARS data. The area with  $T < T_{\text{act}}$  is calculated on seven isentropic surfaces in the lower stratosphere: 390, 410, 430, 460, 490, 520 and 550 K;  $T_{\text{act}}$  on these levels is 197.5, 197.2, 196.8, 196.5, 195.9, 195.3 and 194.5 K, respectively. To get the volume with  $T < T_{\text{act}}$  from 380 through 565 K, the areas at each of the seven levels are multiplied by the estimated altitude associated with that layer and summed. The altitude range associated with each layer is obtained from a standard potential temperature profile as a function of altitude derived from high latitude temperature soundings taken during the 1988–89 through to 2001–02 winters (the same profile was used for  $V_{\text{psc}}$  calculations in refs 13, 16 and 48). These thicknesses are 1.29088, 1.19995, 1.36770, 1.46281, 1.30554, 1.18199 and 1.07382 km for the seven levels listed above. Vortex volume is calculated from vortex area in the same manner. Winter mean  $V_{\text{psc}}$  is calculated over 16 December through to 15 April. Previous studies have shown that  $V_{\text{psc}}$  scaled by the vortex area is a good proxy for chlorine activation and ozone loss potential<sup>17</sup>. Additional temperature and vortex diagnostics are described in Supplementary Information.

**Polar stratospheric cloud and aerosol information.** Particulate backscatter averaged over the polar vortex derived from CALIOP data is used to provide PSC/aerosol information. Total attenuated backscatter at 532 nm,  $b(z)$ , is one of the basic CALIOP Level 1B data products.  $b(z)$  is the sum of the particulate backscatter (due to liquid aerosol and PSCs),  $b_p(z)$ , and molecular backscatter,  $b_m(z)$ .  $b_m(z)$  is calculated using GEOS-5 molecular density profiles (included in the CALIOP Level 1B data files) and a theoretical value for the molecular scattering cross-section<sup>55</sup>. Profiles of  $b_p(z)$  are then produced by subtracting  $b_m(z)$  from  $b(z)$ . Vortex-averaged profiles of  $b_p(z)$  are produced by averaging all CALIOP  $b_p(z)$  profiles located inside the vortex edge (defined using information available in GEOS-5 Derived Meteorological Product (DMP) files for the nearly-coincident Aura MLS data<sup>52</sup>) over the selected time interval.

**MLS trace gas profile measurements and analysis.** Trace gas profile measurements of  $\text{HNO}_3$ , HCl, ClO, ozone and  $\text{N}_2\text{O}$  (a long-lived tracer used to assess descent) are from Aura MLS<sup>45</sup> version 3 retrievals; data quality screening is as recommended in the MLS data quality document<sup>56</sup>. MLS data are retrieved on pressure surfaces; potential temperature as a function of pressure from MLS DMPs<sup>52</sup> calculated from GEOS-5 analyses is used to interpolate to isentropic surfaces. Vortex averages of MLS data are calculated using the  $1.4 \times 10^{-4} \text{ s}^{-1}$  scaled PV contour to define the vortex edge, using PV values from the MLS DMPs<sup>52</sup>. Active chlorine is in the form of ClO mainly during the daytime, and thus measured ClO amounts vary with the solar zenith angle (SZA) at which the measurements are taken. Only daytime ClO measurements are used here. Northern high latitudes are sampled near midday local time, southern high latitudes are sampled in late afternoon, thus the SZA of Aura MLS Antarctic measurements is  $\sim 7^\circ$  higher on average than that in the Arctic. Reactive chlorine partitioning shifts away from ClO at higher SZAs<sup>7,12</sup>, leading to  $\sim 30\%$  lower ClO measured by Aura MLS in the Antarctic than in the Arctic under fully activated

conditions. MLS measurements are unavailable from 27 March through to 20 April 2011 because of an instrument anomaly. Upper Atmosphere Research Satellite (UARS) MLS measurements, used for analysis of 1995–96 and 1996–97, are sparse because of the UARS yaw cycle and other measurement gaps<sup>26</sup>. The time of day of UARS measurements varied through the yaw cycle, in the middle of which no daytime ClO measurements were obtained<sup>10</sup>; thus ClO values shown in 1995–96 and 1996–97 near those dates (including the mid-February 1996 measurements shown in Fig. 2g) are not representative of the degree of chlorine activation.

**Chemical loss calculations.** Chemical ozone loss is quantified by two methods, both widely used for such calculations<sup>7,12,24–28,47,48</sup>. In the 'passive subtraction' method<sup>25–27</sup>, a transport model is used to calculate the evolution of ozone in the absence of chemical changes ('passive' ozone). The difference between passive ozone and observed ozone provides an estimate of chemical loss.

Here, passive ozone is obtained in three different ways. First, MLS observations of  $\text{N}_2\text{O}$ , a long-lived species unaffected by chemical processes, are used to calculate vertical motion, and that estimate of descent is then used to calculate how initial MLS ozone profiles would have evolved in the absence of chemical loss<sup>14</sup>. Second, a 'reverse trajectory' transport model<sup>25,26</sup> is used to transport an initial state based on MLS-observed ozone with no chemistry. Finally, the ATLAS (Alfred Wegener Institute Lagrangian Chemistry/Transport System) chemistry and transport model is run in passive mode<sup>28</sup>, initialized with MLS ozone.

Vortex ozone is also examined in relation to the surfaces on which it is subsiding<sup>12,14,28,48</sup>. The descent rates used here are obtained by averaging radiative heating/cooling rates from the radiation calculation used in the ATLAS model over the polar vortex<sup>48</sup>. These rates are then used to examine vortex-averaged MLS and ozone-sonde data on surfaces of 'spring equivalent potential temperature'<sup>48</sup>, defined as the potential temperature at which air originating at a given level arrived at the end of March. Since the air descended on these surfaces, ozone would have been constant on each such surface in the absence of chemical loss.

The ozone-sonde data used here are all from electrochemical concentration cell (ECC) sondes, made by different manufacturers. Ozone-sonde data quality was assessed in an intercomparison experiment<sup>57</sup> and is discussed in ref. 47. For chemical loss calculations using ozone-sonde data, the profiles are first examined using a procedure for detecting lamination in the profiles; such lamination (an example is shown in Fig. 3f) is associated with mixing in of extra-vortex air, which may obscure the signature of chemical loss. Profiles that have been significantly altered by mixing processes, as indicated by lamination, are excluded from the vortex averages used in the chemical loss calculations. 2010–11 Arctic ozone-sonde data are provided as Supplementary Information.

Results from the ATLAS model passive subtraction calculations, and from the calculations on spring equivalent potential temperature surfaces using the Match network ozone-sonde data, are shown in Fig. 4; all panels show vortex averages. These results have been compared with the results from the other methods described above. While absolute ozone values obtained from different methods/data sets vary significantly (up to  $\sim 0.4 \text{ p.p.m.v.}$  at the end of March 2011), the year-to-year variations in chemical loss calculated using all three methods agree closely, indicating a high degree of precision in the relative amount of calculated loss between different years.

The Alfred Wegener Institute chemical box model, also used as the chemical module in ATLAS, simulates 175 reactions between 48 chemical species in the stratosphere<sup>27,58</sup>. This model was used to perform conceptual runs (Supplementary Fig. 4), started on 1 March with identical initial mixing ratios of all species except  $\text{HNO}_3$  and  $\text{O}_3$ . For these two species values corresponding to 1997 (3 p.p.m.v.  $\text{O}_3$ , 10 p.p.b.v.  $\text{HNO}_3$ ) and 2011 (2.2 p.p.m.v.  $\text{O}_3$ , 6 p.p.b.v.  $\text{HNO}_3$ ) (compare Figs 2a and 4c) were combined to yield four sets of initial conditions. Initial  $\text{ClO}_x$  was 2 p.p.b.v., corresponding to the vortex-averaged  $\text{ClO}_x$  derived by ATLAS from MLS ClO measurements on 1 March 2011. An air parcel at  $70^\circ \text{N}$ , 460 K potential temperature, with a temperature of 193 K throughout March, was used. Heterogeneous reactions took place on liquid aerosols, rather than solid (nitric acid trihydrate, NAT) PSCs, since the widespread existence of the latter is inconsistent with MLS observations of gas-phase  $\text{HNO}_3$  values (Fig. 2a) larger than those the microphysical module predicts if NAT is present. A sensitivity run showed that sporadically occurring solid PSCs did not change the results significantly.

**Column ozone and ozone deficit calculation.** OMI total ozone data were processed with version 8.5 of the TOMS algorithm and have been extensively validated<sup>59</sup>. TOMS data were processed with version 8 of the algorithm. The OMI and TOMS total ozone data used in this study were averaged on a fixed global  $1^\circ \times 1^\circ$  latitude  $\times$  longitude grid. Averages were computed by area-weighting observations based on the overlap of their instantaneous field-of-view with each grid cell. Only data that satisfy quality criteria based on measurement path length and algorithm diagnostic criteria were included in the averaged samples.

Individual total ozone retrievals included in the samples are expected to have a root-mean-squared error of 1–2%.

Total ozone 'deficit' is calculated as the difference between daily values and a reference that is minimally affected by chemical ozone loss. The reference for the Arctic is the daily mean over all Arctic winters from 1978–79 through to 2009–10, from OMI starting in 2004–05 and from TOMS for earlier years<sup>50</sup>. The Antarctic reference state is the daily mean of TOMS measurements for 1979 through to 1981. Because the Antarctic reference state is based on only three years' data for each day, variations in vortex position are not effectively averaged out; this reference is thus less robust than that for the Arctic, so patterns in daily maps may partially reflect differences in vortex position between the reference and the focus day.

51. Manney, G. L., Zurek, R. W., Gelman, M. E., Miller, A. J. & Nagatani, R. The anomalous Arctic lower stratospheric polar vortex of 1992–1993. *Geophys. Res. Lett.* **21**, 2405–2408 (1994).
52. Manney, G. L. *et al.* Solar occultation satellite data and derived meteorological products: Sampling issues and comparisons with Aura MLS. *J. Geophys. Res.* **112**, D24S50, <http://dx.doi.org/10.1029/2007JD008709> (2007).
53. Butchart, N. & Remsberg, E. E. The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface. *J. Atmos. Sci.* **43**, 1319–1339 (1986).
54. Hanson, D. & Mauersberger, K. Laboratory studies of the nitric acid trihydrate: implications for the south polar stratosphere. *Geophys. Res. Lett.* **15**, 855–858 (1988).
55. Hostetler, C. A. *et al.* CALIOP algorithm theoretical basis document. Calibration and Level 1 data products (Technical Report, NASA Langley Research Center, 2006); available at (<http://www-calipso.larc.nasa.gov/resources/pdfs/PC-SCI-201v1.0.pdf>).
56. Livesey, N. J. *et al.* Version 3.3 Level 2 data quality and description document. (Technical Report JPL D-33509, Jet Propulsion Laboratory, 2010); available at ([http://mls.jpl.nasa.gov/data/v3-3\\_data\\_quality\\_document.pdf](http://mls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf)).
57. Smit, H. G. *et al.* Assessment of the performance of ECC-ozone sondes under quasi-flight conditions in the environmental simulation chamber: insights from the Jülich Ozone Sonde Intercomparison Experiment (JOSIE). *J. Geophys. Res.* **112**, D19306, <http://dx.doi.org/10.1029/2006JD007308> (2007).
58. Krämer, M. *et al.* Intercomparison of stratospheric chemistry models under polar vortex conditions. *J. Atmos. Chem.* **45**, 51–77 (2003).
59. McPeters, R. *et al.* Validation of the Aura Ozone Monitoring Instrument total column ozone product. *J. Geophys. Res.* **113**, D15S14, <http://dx.doi.org/10.1029/2007JD008802> (2008).

# A high-resolution map of human evolutionary constraint using 29 mammals

Kerstin Lindblad-Toh<sup>1,2</sup>, Manuel Garber<sup>1\*</sup>, Or Zuk<sup>1\*</sup>, Michael F. Lin<sup>1,3\*</sup>, Brian J. Parker<sup>4\*</sup>, Stefan Washietl<sup>3\*</sup>, Pouya Kheradpour<sup>1,3\*</sup>, Jason Ernst<sup>1,3\*</sup>, Gregory Jordan<sup>5\*</sup>, Evan Mauceli<sup>1\*</sup>, Lucas D. Ward<sup>1,3\*</sup>, Craig B. Lowe<sup>6,7,8\*</sup>, Alisha K. Holloway<sup>9\*</sup>, Michele Clamp<sup>1,10\*</sup>, Sante Gnerre<sup>1\*</sup>, Jessica Alföldi<sup>1</sup>, Kathryn Beal<sup>5</sup>, Jean Chang<sup>1</sup>, Hiram Clawson<sup>6</sup>, James Cuff<sup>11</sup>, Federica Di Palma<sup>1</sup>, Stephen Fitzgerald<sup>5</sup>, Paul Flicek<sup>5</sup>, Mitchell Guttman<sup>1</sup>, Melissa J. Hubisz<sup>12</sup>, David B. Jaffe<sup>1</sup>, Irwin Jungreis<sup>3</sup>, W. James Kent<sup>9</sup>, Dennis Kostka<sup>9</sup>, Marcia Lara<sup>1</sup>, Andre L. Martins<sup>12</sup>, Tim Massingham<sup>5</sup>, Ida Moltke<sup>4</sup>, Brian J. Raney<sup>6</sup>, Matthew D. Rasmussen<sup>3</sup>, Jim Robinson<sup>1</sup>, Alexander Stark<sup>13</sup>, Albert J. Vilella<sup>5</sup>, Jiayu Wen<sup>4</sup>, Xiaohui Xie<sup>1</sup>, Michael C. Zody<sup>1</sup>, Broad Institute Sequencing Platform and Whole Genome Assembly Team†, Kim C. Worley<sup>14</sup>, Christie L. Kovar<sup>14</sup>, Donna M. Muzny<sup>14</sup>, Richard A. Gibbs<sup>14</sup>, Baylor College of Medicine Human Genome Sequencing Center Sequencing Team†, Wesley C. Warren<sup>15</sup>, Elaine R. Mardis<sup>15</sup>, George M. Weinstock<sup>14,15</sup>, Richard K. Wilson<sup>15</sup>, Genome Institute at Washington University†, Ewan Birney<sup>5</sup>, Elliott H. Margulies<sup>16</sup>, Javier Herrero<sup>5</sup>, Eric D. Green<sup>17</sup>, David Haussler<sup>6,8</sup>, Adam Siepel<sup>12</sup>, Nick Goldman<sup>5</sup>, Katherine S. Pollard<sup>9,18</sup>, Jakob S. Pedersen<sup>4,19</sup>, Eric S. Lander<sup>1</sup> & Manolis Kellis<sup>1,3</sup>

**The comparison of related genomes has emerged as a powerful lens for genome interpretation. Here we report the sequencing and comparative analysis of 29 eutherian genomes. We confirm that at least 5.5% of the human genome has undergone purifying selection, and locate constrained elements covering ~4.2% of the genome. We use evolutionary signatures and comparisons with experimental data sets to suggest candidate functions for ~60% of constrained bases. These elements reveal a small number of new coding exons, candidate stop codon readthrough events and over 10,000 regions of overlapping synonymous constraint within protein-coding exons. We find 220 candidate RNA structural families, and nearly a million elements overlapping potential promoter, enhancer and insulator regions. We report specific amino acid residues that have undergone positive selection, 280,000 non-coding elements exapted from mobile elements and more than 1,000 primate- and human-accelerated elements. Overlap with disease-associated variants indicates that our findings will be relevant for studies of human biology, health and disease.**

A key goal in understanding the human genome is to discover and interpret all functional elements encoded within its sequence. Although only ~1.5% of the human genome encodes protein sequence<sup>1</sup>, comparative analysis with the mouse<sup>2</sup>, rat<sup>3</sup> and dog<sup>4</sup> genomes showed that at least 5% is under purifying selection and thus probably functional, of which ~3.5% consists of non-coding elements with probable regulatory roles. Detecting and interpreting these elements is particularly relevant to medicine, as loci identified in genome-wide association studies (GWAS) frequently lie in non-coding sequence<sup>5</sup>.

Although initial comparative mammalian studies could estimate the overall proportion of the genome under evolutionary constraint, they had little power to detect most of the constrained elements—especially the smaller ones. Thus, they focused only on the top 5% of constrained sequence, corresponding to less than ~0.2% of the genome<sup>4,6</sup>. In 2005, we began an effort to generate sequence from a large collection of mammalian genomes with the specific goal of identifying and interpreting functional elements in the human genome on

the basis of their evolutionary signatures<sup>7,8</sup>. Here we report our results to systematically characterize mammalian constraint using 29 eutherian (placental) genomes. We identify 4.2% of the human genome as constrained and ascribe potential function to ~60% of these bases using diverse lines of evidence for protein-coding, RNA, regulatory and chromatin roles, and we present evidence of exaptation and accelerated evolution. All data sets described here are publicly available in a comprehensive data set at the Broad Institute and University of California, Santa Cruz (UCSC).

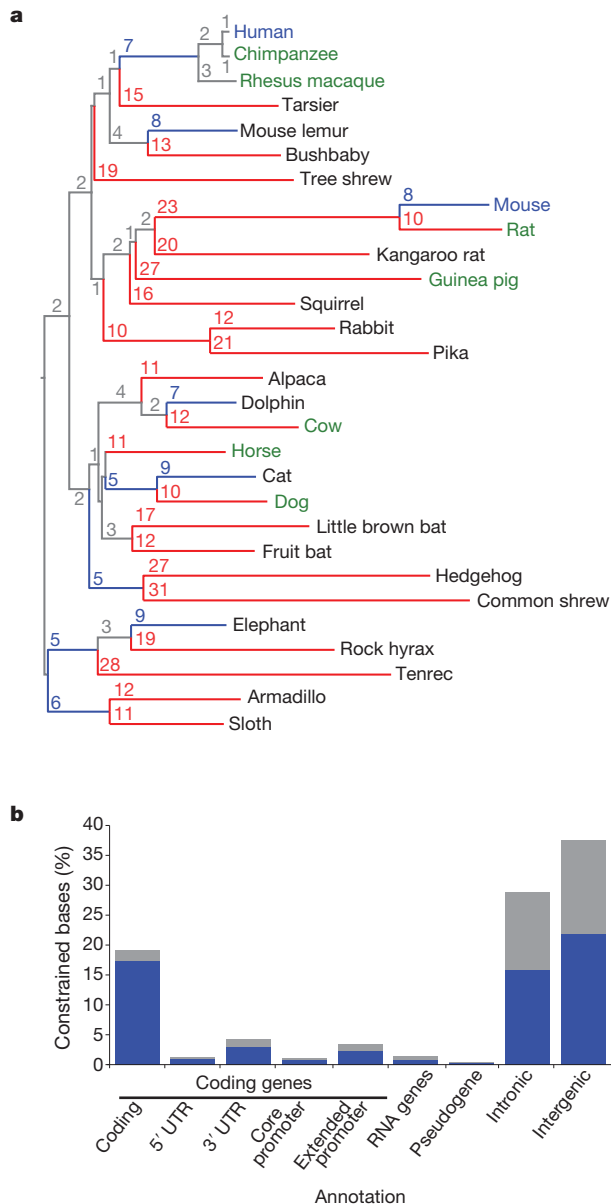
## Sequencing, assembly and alignment

We generated genome sequence assemblies for 29 mammalian species selected to achieve maximum divergence across the four major mammalian clades (Fig. 1a and Supplementary Text 1 and Supplementary Table 1). For nine species, we used genome assemblies based on ~7-fold coverage shotgun sequence, and for 20 species we generated ~2-fold coverage (2×), to maximize the number of species sequenced

<sup>1</sup>Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>2</sup>Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Box 582, SE-751 23 Uppsala, Sweden. <sup>3</sup>MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St. Cambridge, Massachusetts 02139, USA. <sup>4</sup>The Bioinformatics Centre, Department of Biology, University of Copenhagen, DK-2200 Copenhagen, Denmark. <sup>5</sup>EMBL-EBI, Wellcome Trust Genome Campus, Hinxton CB10 1SD, UK. <sup>6</sup>Center for Biomolecular Science and Engineering, University of California, Santa Cruz, California 95064, USA. <sup>7</sup>Department of Developmental Biology, Stanford University, Stanford, California 94305, USA. <sup>8</sup>Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA. <sup>9</sup>Gladstone Institutes, University of California, 1650 Owens Street, San Francisco, California 94158, USA. <sup>10</sup>BioTeam Inc, 7 Derosier Drive, Middleton, Massachusetts 01949, USA. <sup>11</sup>Research Computing, Division of Science, Faculty of Arts and Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>12</sup>Department of Biological Statistics & Computational Biology, Cornell University, Ithaca, New York 14853, USA. <sup>13</sup>Research Institute of Molecular Pathology (IMP), A-1030 Vienna, Austria. <sup>14</sup>Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA. <sup>15</sup>Genome Institute at Washington University, Washington University School of Medicine, 4444 Forest Park Blvd., Saint Louis, Missouri 63108, USA. <sup>16</sup>Genome Informatics Section, Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892 USA. <sup>17</sup>NISC Comparative Sequencing Program, Genome Technology Branch and NIH Intramural Sequencing Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892 USA. <sup>18</sup>Institute for Human Genetics, and Division of Biostatistics, University of California, 1650 Owens Street, San Francisco, California 94158, USA. <sup>19</sup>Department of Molecular Medicine (MOMA), Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark.

\*These authors contributed equally to this work.

†A full list of authors and their affiliations appears at the end of paper.



**Figure 1 | Phylogeny and constrained elements from the 29 eutherian mammalian genome sequences.** **a**, A phylogenetic tree of all 29 mammals used in this analysis based on the substitution rates in the MultiZ alignments. Organisms with finished genome sequences are indicated in blue, high quality drafts in green and 2x assemblies in black. Substitutions per 100 bp are given for each branch; branches with  $\geq 10$  substitutions are coloured red, blue indicates  $< 10$  substitutions. **b**, At 10% FDR, 3.6 million constrained elements can be detected encompassing 4.2% of the genome, including a substantial fraction of newly detected bases (blue) compared to the union of the HMRD 50-bp + Siepel vertebrate elements<sup>17</sup> (see Supplementary Fig. 4b for comparison to HMRD elements only). The largest fraction of constraint can be seen in coding exons, introns and intergenic regions. For unique counts, the analysis was performed hierarchically: coding exons, 5' UTRs, 3' UTRs, promoters, pseudogenes, non-coding RNAs, introns, intergenic. The constrained bases are particularly enriched in coding transcripts and their promoters (Supplementary Fig. 4c).

with available resources on capillary machines. Twenty genomes are first reported here, and nine were previously described (see Supplementary Information).

The power to detect constrained elements depends largely on the total branch length of the phylogenetic tree connecting the species<sup>9</sup>. The 29 mammals correspond to a total effective branch length of  $\sim 4.5$

substitutions per site, compared to  $\sim 0.68$  for the human–mouse–rat–dog comparison (HMRD), and thus should offer greater power to detect evolutionary constraint: the probability that a genomic sequence not under purifying selection will remain fixed across all 29 species is  $P_1 < 0.02$  for single bases and  $P_{12} < 10^{-25}$  for 12-nucleotide sequences, compared to  $P_1 \sim 0.50$  and  $P_{12} \sim 10^{-3}$  for HMRD.

For mammals for which we generated 2x coverage, our assisted assembly approach<sup>10</sup> resulted in a typical contig size  $N50_C$  of 2.8 kb and a typical scaffold size  $N50_S$  of 51.8 kb (Supplementary Text 2 and Supplementary Table 1) and high sequence accuracy (96% of bases had quality score Q20, corresponding to a  $< 1\%$  error rate)<sup>11</sup>. Compared to high-quality sequence across the 30 Mb of the ENCODE pilot project<sup>12</sup>, we estimated average error rates of 1–3 miscalled bases per kilobase<sup>11</sup>, which is  $\sim 50$ -fold lower than the typical nucleotide sequence difference between the species, enabling high-confidence detection of evolutionary constraint (Supplementary Text 3).

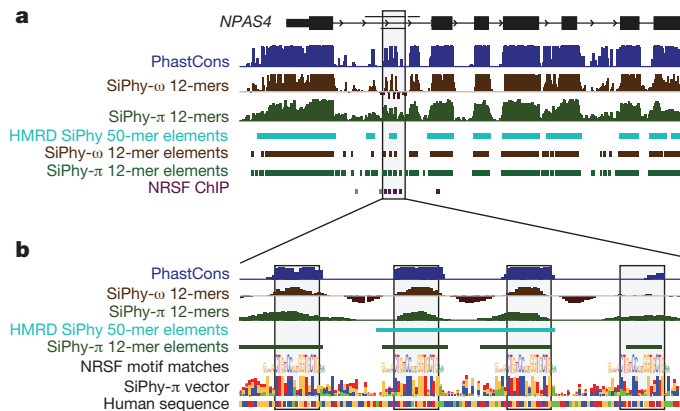
We based our analysis on whole-genome alignments by MultiZ (Supplementary Text 4). The average number of aligned species was 20.9 at protein-coding positions in the human genome and 23.9 at the top 5% HMRD-conserved non-coding positions, with an average branch length of 4.3 substitutions per base in these regions (Supplementary Figs 1 and 2). In contrast, whole-genome average alignment depth is only 17.1 species with 2.9 substitutions per site, probably due to large deletions in non-functional regions<sup>4</sup>. The depth at ancestral repeats is 11.4 (Supplementary Fig. 1a), consistent with repeats being largely non-functional<sup>2,4</sup>.

## Detection of constrained sequence

Our analysis did not substantially change the estimate of the proportion of genome under selection. By comparing genome-wide conservation to that of ancestral repeats, we estimated the overall fraction of the genome under evolutionary constraint to be 5.36% at 50-bp windows (5.44% at 12-bp windows), using the SiPhy- $\omega$  statistic<sup>13</sup>, a measure of overall substitution rate (Supplementary Fig. 3), consistent with previous similar estimates<sup>2,4,14</sup>. However, alternative methods<sup>15,16</sup> and different ways of correcting for the varying alignment depths give higher estimates (see Supplementary Text 5 for details).

The additional species had a marked effect on our ability to identify the specific elements under constraint. With 29 mammals, we pinpoint 3.6 million elements spanning 4.2% of the genome, at a finer resolution of 12 bp (Fig. 1b and Supplementary Text 6, Supplementary Fig. 4, Supplementary Tables 2 and 3), compared to  $< 0.1\%$  of the genome for HMRD 12-bp elements and 2.0% for HMRD 50-bp elements<sup>4</sup>. Elements previously detected using five vertebrates<sup>17</sup> also detect a larger fraction of the genome ( $\sim 4.1\%$ ), but only cover 45% of the mammalian elements detected here, suggesting that a large fraction of our elements are mammalian specific. The mean element size (36 bp) is considerably shorter than both previously detected HMRD elements (123 bp) and five-vertebrate elements (104 bp)<sup>17</sup>. For example, it is now possible to detect individual binding sites for the neuron-restrictive silencer factor (NRSF) in the promoter of the *NPAS4* gene, which are beyond detection power in previous data sets (Fig. 2 and Supplementary Fig. 5). We found a similar regional distribution of 12-bp elements (including the 2.6 million newly detected constrained elements) to previously detected HMRD elements ( $r = 0.94$ , Supplementary Fig. 6). Similar results were obtained with the PhastCons<sup>17</sup> statistic (see Supplementary Text 6).

Using a new method, SiPhy- $\pi$ , sensitive not just to the substitution rate but also to biases in the substitution pattern (for example, positions free to mutate between G and T only, Supplementary Fig. 7), we detected an additional 1.3% of the human genome in constrained elements (see Supplementary Tables 2 and 3). Most of the newly detected constrained nucleotides extend elements found by rate-based methods, but 22% of nucleotides lie in new elements (average length 17 bp) and are enriched in non-coding regions.



**Figure 2 | Identification of four NRSF-binding sites in *NPAS4*.** **a**, The neurological gene *NPAS4* has many constrained elements overlapping introns and the upstream intergenic region. The grey shaded box contained only one constrained element using HMRD, whereas analysis of 29 mammalian sequences reveals four smaller elements. **b**, These four constrained elements in the first intron correspond to binding sites for the NRSF transcription factor, known to regulate neuronal lineages.

### Constraint within the human population

We observed that the evolutionary constraint acting on the 29 mammals is correlated with constraint within the human population, as assessed from human polymorphism data (Supplementary Text 7) and consistent with previous studies<sup>18</sup>. Mammalian constrained elements show a depletion in single-nucleotide polymorphisms (SNPs)<sup>19</sup>, and more constrained elements show even greater depletion. For example, in the top 1% most strongly conserved non-coding regions, SNPs occur at a 1.9-fold lower rate than the genome average, and the derived alleles have a lower frequency, consistent with purifying selection at many of these sites in the human population.

Moreover, at positions with biased substitution patterns across mammals, the observed human SNPs show a similar bias to the one observed across mammals (Supplementary Fig. 7). Thus, not only are constrained regions less likely to exhibit polymorphism in humans, but when such polymorphisms are observed, the derived alleles in humans tend to match the alleles present in non-human mammals, indicating a preference for the same alleles across both mammalian and human evolution.

### Functional annotation of constraint

We first studied the overlap of the 3.6 million evolutionarily constrained elements ( $\omega < 0.8$  and  $P < 10^{-15}$ ) with known gene annotations (Fig. 1b). Roughly 30% of constrained elements were associated with protein-coding transcripts: ~25.3% overlap mature messenger RNAs (including 19.6% in coding exons, 1.2% in 5' untranslated regions (5' UTRs) and 4.4% in 3' UTRs), and an additional 4.4% reside within 2 kb of transcriptional start sites (1.2% reside within 200 bases).

The majority of constrained elements, however, reside in intronic and intergenic regions (29.7% and 38.6%, respectively). To study their biological roles and provide potential starting points to understand these large and mostly uncharted territories, we next studied their overlap with evolutionary signatures<sup>7,8,20,21</sup> characteristic of specific types of features and a growing collection of public large-scale experimental data.

### Protein-coding genes and exons

Despite intense efforts to annotate protein-coding genes over the past decade<sup>20,22–24</sup>, we detected 3,788 candidate new exons (a 2% increase) using evolutionary signatures characteristic of protein-coding exons<sup>25</sup>. Of these, 54% reside outside transcripts of protein-coding genes, 19% within introns, and 13% in UTRs of known coding genes (Supplementary Text 8, Supplementary Tables 4 and 5). Our methods recovered

92% of known coding exons that were larger than 10 codons and fall in syntenic regions, the remainder showing non-consensus splice sites, unusual features, or poor conservation.

The majority of new exon candidates (>58%) are supported by evidence of transcription measured in 16 human tissues<sup>26</sup> (Supplementary Fig. 8a) or similarity to known Pfam protein domains. Thirty-one per cent of intronic and 13% of intergenic predictions extend known transcripts, and 5% and 11% respectively reside in new transcript models. The newly detected exons are more tissue specific than known exons (mean of 3 tissues versus 12) and are expressed at fivefold lower levels. Directed experiments and manual curation will be required to complete the annotation of the few hundred protein-coding genes that probably remain unannotated<sup>27</sup>.

We found apparent stop codon readthrough<sup>28</sup> of four genes based on continued protein-coding constraint after an initial conserved stop codon<sup>29</sup> and until a subsequent stop codon (Supplementary Text 9 and Supplementary Fig. 8b). Readthrough in *SACMIL* could be triggered by an 80-base conserved RNA stem loop predicted by RNAz<sup>30</sup>, lying four bases downstream of the readthrough stop codon.

We also detected coding regions with a very low synonymous substitution rate, indicating additional sequence constraints beyond the amino acid level (Supplementary Text 9). We found >10,000 such synonymous constraint elements (SCEs) in more than one-quarter of all human genes<sup>31</sup>. Initial analysis indicates potential roles in splicing regulation (34% span an exon–exon junction), A-to-I editing, microRNA (miRNA) targeting and developmental regulation. *HOX* genes contain several top candidates (Fig. 3a), including two previously validated developmental enhancers<sup>32,33</sup>.

### RNA structures and families of structural elements

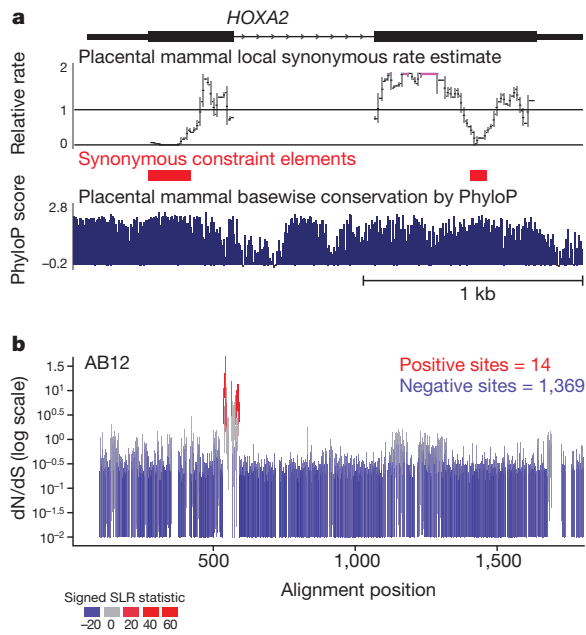
We next used evolutionary signatures characteristic of conserved RNA secondary structures<sup>34</sup> to reveal 37,381 candidate structural elements (Supplementary Text 10 and Supplementary Fig. 9a), covering ~1% of constrained regions. For example, the *XIST* large intergenic non-coding RNA (lincRNA), known to bind chromatin and enable X-chromosome inactivation<sup>35</sup>, contains a newly predicted structure in its 3' end (Supplementary Fig. 9b, f)—distinct from other known structures<sup>36</sup>—which seems to be the source of chromatin-associated short RNAs<sup>37</sup>.

Sequence- and structure-based clustering of predictions outside protein-coding exons revealed 1,192 novel families of structural RNAs (Supplementary Text 10). We focused on a high-scoring subset consisting of 220 families with 725 instances, which also showed the highest thermodynamic stability<sup>30</sup> (Supplementary Figs 9a and 10), DNase hypersensitivity, expression pattern correlation across tissues and intergenic expression enrichment (Supplementary Fig. 9a). We also expanded both known and novel families by including additional members detected by homology to existing members.

Noteworthy examples include: a glycyl-tRNA family, including a new member in *POPI*, involved in tRNA maturation and probably involved in feedback regulation of *POPI*; three intronic families of long hairpins in ion-channel genes known to undergo A-to-I RNA editing and possibly involved in regulation of the editing event; an additional member of a family of 5' UTR hairpins overlapping the start codon of collagen genes and potential new miRNA genes that extend existing families<sup>37</sup>.

Two of the largest novel families consist of short AU-rich hairpins of 6–7 bp that share the same strong consensus motif in their stem. These occur in the 3' UTRs of genes in several inflammatory response pathways, the post-transcriptional regulation of which often involves AU-rich elements (AREs). Indeed, two homologous hairpins in *TNF* and *CSF3* correspond to known mRNA-destabilization elements, suggesting roles in mRNA stability for the two families<sup>37</sup>.

Lastly, a family of six conserved hairpin structures (Supplementary Fig. 9g) was found in the 3' UTR of the *MAT2A* gene<sup>37</sup>, which is involved in the synthesis of S-adenosylmethionine (SAM), the primary



**Figure 3 | Examination of evolutionary signatures identifies SCEs and evidence of positive selection.** **a**, Two regions within the *HOXA2* open reading frame are identified as SCEs (red), corresponding to overlapping functional elements within coding regions. Note that the synonymous rate reductions are not obvious from the base-wise conservation measure (in blue). Both elements have been characterized as enhancers driving *HOXA2* expression in distinct segments of the developing mouse hindbrain. The element in the first exon encodes Hox-Pbx-binding sites and drives expression in rhombomere 4 (ref. 33), whereas the element in the second exon contains Sox-binding sites and drives expression in rhombomere 2 (ref. 32). Synonymous constraint elements are also found in most other *HOX* genes, and up to a quarter of all genes. **b**, Although ~85% of genes show only negative (purifying) selection and 9% of genes show uniform positive selection, the remaining 6% of genes, including *AB12*, show only localized regions of positively selected sites. Each vertical bar covers the estimated 95% confidence interval for dN/dS at that site (with values of 0 truncated to 0.01 to accommodate the log scaling), and bars are coloured according to a signed version of the SLR statistic for non-neutral evolution: blue for sites under purifying selection, grey for neutral sites and red for sites under positive selection.

methyl donor in human cells. All six hairpins consist of a 12–18-bp stem and a 14-bp loop region with a deeply conserved sequence motif (Supplementary Fig. 9e), and may be involved in sensing SAM concentrations, which are known to affect *MAT2A* mRNA stability<sup>38</sup>.

### Conservation patterns in promoters

As different types of conservation in promoters may imply distinct biological functions<sup>39</sup>, we classified the patterns of conservation within core promoters into three categories: (1) those with uniformly 'high' constraint (7,635 genes, 13,996 transcripts); (2) uniformly 'low' constraint (2,879 genes, 4,135 transcripts); and (3) 'intermittent' constraint, consisting of alternating peaks and troughs of conservation (14,271 genes and 29,814 transcripts) (Supplementary Fig. 11a). High and intermittent constraint promoters are both associated with CpG islands (~66%), whereas low constraint promoters have significantly lower overlap (~41%), and all three classes show similar overlap with functional TATA boxes (2–3%, see Supplementary Text 11).

These groups show distinct Gene Ontology enrichments (Supplementary Fig. 11b), with high-constraint promoters involved in development ( $P$  with Bonferroni correction ( $P_{\text{Bonf}} < 10^{-30}$ ), intermittent constraint in basic cellular functions ( $P_{\text{Bonf}} < 5 \times 10^{-4}$ ), and low-constraint promoters in immunity, reproduction and perception, functions expected to be under positive selection and lineage-specific adaptation<sup>2</sup>.

High constraint may reflect cooperative binding of many densely binding factors, as previously suggested for developmental genes<sup>6</sup>. Intermittent constraint promoters, the peak-spacing distribution of which was suggestive of the periodicity of the DNA helix turns, may reflect loosely interacting factors (Supplementary Fig. 11c, d). Low constraint may reflect rapid motif turnover, under neutral drift or positive selection.

### Identifying specific instances of regulatory motifs

Data from just four species (HMRD) was sufficient to create a catalogue of known and novel motifs with many conserved instances across the genome<sup>21</sup>. The power to discover such motifs was high, because one can aggregate data across hundreds of motif instances. Not surprisingly, the additional genomes therefore had little effect on the ability to discover new motifs (known motifs showed 99% correlation in genome-wide motif conservation scores, Supplementary Figs 12 and 13).

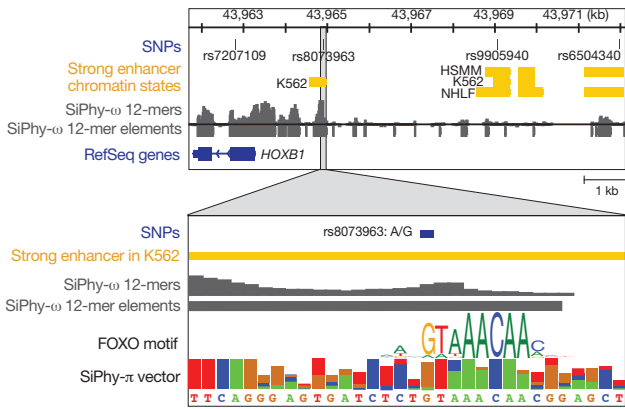
In contrast, the 29 mammalian genomes markedly improved our ability to detect individual motif instances, making it possible to predict specific target sites for 688 regulatory motifs corresponding to 345 transcription factors (Supplementary Fig. 14). We chose to identify motif instances at a false discovery rate (FDR) of 60%, representing a reasonable compromise between specificity and sensitivity given the available discovery power (Supplementary Text 12), and matching the experimental specificity of chromatin immunoprecipitation (ChIP) experiments for identifying biologically significant targets<sup>40</sup>. Higher levels of stringency could be obtained by sequencing additional species.

We identified 2.7 million conserved instances (Supplementary Table 6), enabling the construction of a regulatory network linking 375 motifs to predicted targets, with a median of 21 predicted regulators per target gene (25th percentile, 10; 75th percentile, 39). The number of target sites (average, 4,277; 25th percentile, 1,407; 75th percentile, 10,782) are comparable to those found in ChIP experiments, and have the advantage that they are detected at nucleotide resolution, enabling us to use them to interpret disease-associated variants for potential regulatory functions. However, some motifs never reached high confidence values, and others did so at very few instances.

The motif-based targets show strong agreement with experimentally defined binding sites from ChIP experiments (Supplementary Table 7). For long and distinct motifs, such as CTCF and NRSF, the fraction of instances overlapping experimentally observed binding matches the fraction predicted by the confidence score (for example, at 80% confidence 70% of NRSF motif instances overlapped bound sites, and at ~50% confidence 40% overlapped), despite potential confounding aspects such as condition-specific binding, overlapping motifs between factors, or non-specific binding. Moreover, increasing confidence levels showed increasing overlap with experimental binding (Supplementary Figs 14–16). For example, *YY1* enrichment for bound sites increased from 42-fold to 168-fold by focusing on conserved instances. Lastly, combining motif conservation and experimental binding led to increased enrichment for candidate tissue-specific enhancers, suggesting that the two provide complementary information. Within bound regions, the evolutionary signal reveals specific motif instances with high precision (for example, Figs 2 and 4 and Supplementary Fig. 17).

### Chromatin signatures

To suggest potential functions for the ~68% of 'unexplained' constrained elements outside coding regions, UTRs or proximal promoters, we used chromatin state maps from CD4 T cells<sup>41</sup> (Supplementary Fig. 18) and nine diverse cell types<sup>42</sup> (Supplementary Text 13 and Supplementary Fig. 19). In T cells, constrained elements were most enriched for promoter-associated states (up to fivefold), an insulator state and a specific repressed state (2.2-fold), and numerous enhancer



**Figure 4 | Using constraint to identify candidate mutations.** Conservation can help us resolve amid multiple SNPs the ones that disrupt conserved functional elements and are likely to have regulatory roles. In this example, a SNP (rs6504340) associated with tooth development is strongly linked to a conserved intergenic SNP, rs8073963, 7.1 kb away, which disrupts a deeply conserved Forkhead-family motif in a strong enhancer. Although the SNPs shown here stem from GWAS on HapMap data, the same principle should be applicable to associated variants detected by resequencing the region of interest.

states (1.5–2-fold), together covering 7.1% of the unexplained elements at 2.1-fold enrichment. In the nine cell types, enriched promoter, enhancer and insulator states cover 36% of unexplained elements at ~1.75-fold enrichment, with locations active in multiple cell types showing even stronger enrichment (Supplementary Fig. 20).

Overall, chromatin states indicate possible functions (at 1.74-fold enrichment) for 37.5% ( $N = 987,985$ ) of unexplained conserved elements (27% of all conserved elements), suggesting meaningful association for at least 16% of unexplained constrained bases. Although current experiments only provide nucleosome-scale (~200-bp) resolution, we expect higher-resolution experimental assays that more precisely pinpoint regulatory regions to show further increases in enrichment. The increased overlap observed with additional cell types suggests that new cell types will help elucidate additional elements. Of course, further experimental tests will be required to validate the predicted functional roles.

### Accounting for constrained elements

Overall, ~30% of constrained elements overlap were associated with protein-coding transcripts, ~27% overlap specific enriched chromatin states, ~1.5% novel RNA structures, and ~3% conserved regulatory motif instances (Supplementary Text 13, 14). Together, ~60% of constrained elements overlap one of these features, with enrichments ranging from 1.75-fold for chromatin states (compared to unannotated regions) up to 17-fold for protein-coding exons (compared to the whole genome).

### Implications for interpreting disease-associated variants

In the non-protein-coding genome, SNPs associated with human diseases in genome-wide association studies are 1.37-fold enriched for constrained regions, relative to HapMap SNPs (Supplementary Text 15 and Supplementary Table 8). This is notable because only a small proportion of the associated SNPs are likely to be causative, whereas the rest are merely in linkage disequilibrium with causative variants.

Accordingly, constrained elements should be valuable in focusing the search for causative variants among multiple variants in linkage disequilibrium. For example, in an intergenic region between *HOXB1* and *HOXB2* associated with tooth development phenotypes<sup>43</sup>, the reported SNP (rs6504340) is not conserved, but a linked SNP (rs8073963) sits in a constrained element 7.1 kb away. Moreover, rs8073963 disrupts a deeply conserved FOXO2 motif instance within a predicted enhancer (Fig. 4), making it a candidate mutation for

further follow-up. Similar examples of candidate causal variants are found for diverse phenotypes such as height or multiple sclerosis, and similar analyses could be applied to case-control resequencing data.

### Evolution of constrained elements

We next sought to identify signatures of positive selection that may accompany functional adaptations of different species to diverse environments and new ecosystems.

### Codon-specific selection

We used the ratio dN/dS of non-synonymous to synonymous codon substitutions as evidence of positive selection ( $>1$ ) or negative selection ( $<1$ ). Although dN/dS is typically calculated for whole genes, the additional mammals sequenced enabled analysis at the codon level: simulations predicted a 250-fold gain in sensitivity compared to HMRD, identifying 53% of positive sites at 5% FDR (Supplementary Text 16).

Applying this test to 6.05 million codons in 12,871 gene trees, we found evidence of strong purifying selection ( $dN/dS < 0.5$ ) for 84.2% of codons and positive selection ( $dN/dS > 1.5$ ) for 2.4% of codons (with 94.1% of sites  $<1$  and 5.9%  $>1$ ; Supplementary Table 9). At 5% FDR, we found 15,383 positively selected sites in 4,431 proteins. The genes fell into three classes based on the distribution of selective constraint: 84.8% of genes show uniformly high purifying selection, 8.9% show distributed positive selection across their length and 6.3% show localized positive selection concentrated in small clusters (Fig. 3b and Supplementary Fig. 21, Supplementary Tables 10 and 11).

Genes with distributed positive selection were enriched in such functional categories as immune response ( $P_{\text{Bonf}} < 10^{-16}$ ) and taste perception ( $P_{\text{Bonf}} < 10^{-10}$ ), which are known to evolve rapidly, but also in some unexpected functions such as meiotic chromosome segregation ( $P_{\text{Bonf}} < 10^{-23}$ ) and DNA-dependent regulation of transcription ( $P_{\text{Bonf}} < 10^{-19}$ ; Supplementary Table 12). Localized positive selection was enriched in core biochemical processes, including microtubule-based movement ( $P_{\text{Bonf}} < 10^{-10}$ ), DNA topological change ( $P_{\text{Bonf}} < 10^{-4}$ ) and telomere maintenance ( $P_{\text{Bonf}} < 7 \times 10^{-3}$ ), suggesting adaptation at important functional sites.

Focusing on 451 unique Pfam protein-domain annotations, we found abundant purifying selection, with 225 domains showing purifying selection for >75% of their sites, and 447 domains showing negative selection for >50% of their sites (Supplementary Table 13). Domains with substantial fractions of positively selected sites include CRAL/TRIO involved in retinal binding (2.6%), proteinase-inhibitor-cystatin involved in bone remodelling (2.2%) and the secretion-related EMP24/GOLD/P24 family (1.6%).

### Exaptation of mobile elements

Mobile elements provide an elegant mechanism for distributing a common sequence across the genome, which can then be retained in locations where it confers advantageous regulatory functions to the host—a process termed exaptation. Our data revealed >280,000 mobile element exaptations common to mammalian genomes covering ~7 Mb (Supplementary Text 17), a considerable expansion from the ~10,000 previously recognized cases<sup>44</sup>. Of the ~1.1 million constrained elements that arose during the 90 million years between the divergence from marsupials and the eutherian radiation, we can trace >19% to mobile element exaptations. Often only a small fraction (median ~11%) of each mobile element is constrained, in some cases matching known regulatory motifs. Recent exaptations are generally found near ancestral regulatory elements, except in gene deserts, which are abundant in ancestral elements but show few recent exaptations ( $P < 10^{-300}$ , Supplementary Fig. 22).

### Accelerated evolution in the primate lineage

Lineage-specific rapid evolution in ancestrally constrained elements previously revealed human positive selection associated with brain

and limb development<sup>45</sup>. Applying this signature to the human and primate lineages, we identified 563 human-accelerated regions (HARs) and 577 primate-accelerated regions (PARs) at FDR <10% (Supplementary Text 18, Supplementary Tables 14 and 15), significantly expanding the 202 previously known HARs<sup>46</sup>. Fifty-four HARs (9.4%) and 49 PARs (8.5%) overlap enhancer-associated chromatin marks and experimentally validated enhancers (Supplementary Text 18). Substitution patterns in HARs suggest that GC-biased gene conversion (BGC) is not responsible for the accelerated evolution in the vast majority of these regions (~15% show evidence of BGC).

Genes harbouring or neighbouring HARs and PARs are enriched for extracellular signalling, receptor activity, immunity, axon guidance, cartilage development and embryonic pattern specification (Supplementary Fig. 23). For example, the *FGF13* locus associated with an X-linked form of mental retardation contains four HARs near the 5' ends of alternatively spliced isoforms of *FGF13* expressed in the nervous system, epithelial tissues and tumours, suggesting human-specific changes in isoform regulation (Supplementary Fig. 24).

## Discussion

Comparative analysis of 29 mammalian genomes reveals a high-resolution map of >3.5 million constrained elements that encompass ~4% of the human genome and suggest potential functional classes for ~60% of the constrained bases; the remaining 40% show no overlap and remain uncharacterized. We report previously undetected exons and overlapping functional elements within protein-coding sequence, new classes of RNA structures, promoter conservation profiles and predicted targets of transcriptional regulators. We also provide evidence of evolutionary innovation, including codon-specific positive selection, mobile element exaptation and accelerated evolution in the primate and human lineages.

By focusing our comparison on only eutherian mammals, we discover functional elements relevant to this clade, including recent eutherian innovations. This is especially important for discovering regulatory elements, which can be subject to rapid turnover<sup>47</sup>. Indeed, a previous comparison indicated that only 80% of 50-bp non-coding elements are shared with opossum<sup>48</sup>, and the current 12-bp analysis shows ~64% of non-coding elements shared with opossum, and only 6% with stickleback fish. Many eutherian elements are thus probably missing from previous maps of vertebrate constraint<sup>17</sup>.

Sequencing of additional species should enable discovery of lineage-specific elements within mammalian clades, and provide increased resolution for shared mammalian constraint. We estimate that 100–200 eutherian mammals (15–25 neutral substitutions per site) will enable single-nucleotide resolution. The majority of this branch length is present within the Laurasiatherian and Euarchontoglires branches, which also contain multiple model organisms. These are ideal next targets for sequencing as part of the Genome 10K effort<sup>49</sup>, aiming to sequence 10,000 vertebrate species. Within the primate clade, a branch length of ~1.5 could probably be achieved, enabling primate-specific selection studies, albeit at lower resolution. Lastly, human-specific selection should be detectable by combining data across genomic regions and by comparing thousands of humans<sup>50</sup>.

The constrained elements reported here can be used to prioritize disease-associated variants for subsequent study, providing a powerful lens for elucidating functional elements in the human genome complementary to ongoing large-scale experimental endeavours such as ENCODE and Roadmap Epigenomics. Experimental studies require prior knowledge of the biochemical activity sought and reveal regions active in specific cell types and conditions. Comparative approaches provide an unbiased catalogue of shared functional regions independent of biochemical activity or condition, and thus can capture experimentally intractable or rare activity patterns. With increasing branch length, they can provide information on ancestral and recent selective pressures across mammalian clades and within

the human population. Ultimately, the combination of disease genetics, comparative and population genomics and biochemical studies have important implications for understanding human biology, health and disease.

## METHODS SUMMARY

A full description of materials and methods, including sample selection and sequencing strategy, assembly strategies and results, error estimation and correction, alignment details, estimation of genome portion under constraint, detection of constrained elements, mammalian constraint versus human polymorphism, protein coding genes, detection of stop codon readthrough and synonymous constraint elements, RNA structure detection, patterns of promoter constraint, regulatory motif discovery, correlation with chromatin state information, overall accounting of constraint elements, comparison with disease-associated variants, detection of codon-specific positive selection, exaptation of ancestral repeat elements, and human and primate accelerated regions is available in Supplementary Information. All animal experiments were approved by the MIT Committee for Animal Care.

Received 25 January; accepted 5 September 2011.

Published online 12 October 2011.

1. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
2. Waterston, R. H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
3. Gibbs, R. A. *et al.* Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493–521 (2004).
4. Lindblad-Toh, K. *et al.* Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).
5. Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science* **322**, 881–888 (2008).
6. Bejerano, G. *et al.* Ultraconserved elements in the human genome. *Science* **304**, 1321–1325 (2004).
7. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).
8. Stark, A. *et al.* Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* **450**, 219–232 (2007).
9. Cooper, G. M., Brudno, M., Green, E. D., Batzoglou, S. & Sidow, A. Quantitative estimates of sequence divergence for comparative analyses of mammalian genomes. *Genome Res.* **13**, 813–820 (2003).
10. Gnerre, S., Lander, E. S., Lindblad-Toh, K. & Jaffe, D. B. Assisted assembly: how to improve a *de novo* genome assembly by using related species. *Genome Biol.* **10**, R88 (2009).
11. Hubisz, M. J., Lin, M. F., Kellis, M. & Siepel, A. Error and error mitigation in low-coverage genome assemblies. *PLoS ONE* **14**, e17034 (2011).
12. Thomas, J. W. *et al.* Comparative analyses of multi-species sequences from targeted genomic regions. *Nature* **424**, 788–793 (2003).
13. Garber, M. *et al.* Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics* **25**, i54–i62 (2009).
14. Chiaromonte, F. *et al.* The share of human genomic DNA under selection estimated from human-mouse genomic alignments. *Cold Spring Harb. Symp. Quant. Biol.* **68**, 245–254 (2003).
15. Davydov, E. V. *et al.* Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLOS Comput. Biol.* **6**, e1001025 (2010).
16. Meader, S., Ponting, C. P. & Lunter, G. Massive turnover of functional sequence in human and other mammalian genomes. *Genome Res.* **20**, 1335–1343 (2010).
17. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
18. Drake, J. A. *et al.* Conserved noncoding sequences are selectively constrained and not mutation cold spots. *Nature Genet.* **38**, 223–227 (2006).
19. Keinan, A., Mullikin, J. C., Patterson, N. & Reich, D. Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nature Genet.* **39**, 1251–1255 (2007).
20. Clamp, M. *et al.* Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl Acad. Sci. USA* **104**, 19428–19433 (2007).
21. Xie, X. *et al.* Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338–345 (2005).
22. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
23. Siepel, A. *et al.* Targeted discovery of novel human exons by comparative genomics. *Genome Res.* **17**, 1763–1773 (2007).
24. Pruitt, K. D. *et al.* The consensus coding sequence (CCDS) project: identifying a common protein-coding gene set for the human and mouse genomes. *Genome Res.* **19**, 1316–1323 (2009).
25. Lin, M. F., Jungreis, I. & Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein-coding and non-coding regions. *Bioinformatics* **27**, i275–i282 (2010).
26. Guttman, M. *et al.* *Ab initio* reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature Biotechnol.* **28**, 503–510 (2010).

27. Harrow, J. *et al.* GENCODE: producing a reference annotation for ENCODE. *Genome Biol.* **7** (suppl. 1), 1–9 (2006).
28. Lin, M. F. *et al.* Revisiting the protein-coding gene catalog of *Drosophila melanogaster* using 12 fly genomes. *Genome Res.* **17**, 1823–1836 (2007).
29. Jungreis, I. *et al.* Evidence of abundant stop codon readthrough in *Drosophila* and other metazoa. *Genome Res.* doi:10.1101/gr.119974.110 (in the press).
30. Washietl, S., Hofacker, I. L. & Stadler, P. F. Fast and reliable prediction of noncoding RNAs. *Proc. Natl Acad. Sci. USA* **102**, 2454–2459 (2005).
31. Lin, M. F. *et al.* Locating protein-coding sequences under selection for additional, overlapping functions in 29 mammalian genomes. *Genome Res.* doi:10.1101/gr.108753.110 (in the press).
32. Tumpel, S., Cambrono, F., Sims, C., Krumlauf, R. & Wiedemann, L. M. A regulatory module embedded in the coding region of *Hoxa2* controls expression in rhombomere 2. *Proc. Natl Acad. Sci. USA* **105**, 20077–20082 (2008).
33. Lampe, X. *et al.* An ultraconserved Hox–Pbx responsive element resides in the coding sequence of *Hoxa2* and is active in rhombomere 4. *Nucleic Acids Res.* **36**, 3214–3225 (2008).
34. Pedersen, J. S. *et al.* Identification and classification of conserved RNA secondary structures in the human genome. *PLOS Comput. Biol.* **2**, e33 (2006).
35. Lee, J. T. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev.* **23**, 1831–1842 (2009).
36. Maenner, S. *et al.* 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol.* **8**, e1000276 (2010).
37. Parker, B. J. *et al.* New families of human regulatory RNA structures identified by comparative analysis of vertebrate genomes. *Genome Res.* doi:10.1101/gr.112516.110 (in the press).
38. Martinez-Chantar, M. L. *et al.* L-methionine availability regulates expression of the methionine adenosyltransferase 2A gene in human hepatocarcinoma cells: role of S-adenosylmethionine. *J. Biol. Chem.* **278**, 19885–19890 (2003).
39. Baek, D., Davis, C., Ewing, B., Gordon, D. & Green, P. Characterization and predictive discovery of evolutionarily conserved mammalian alternative promoters. *Genome Res.* **17**, 145–155 (2007).
40. Kheradpour, P., Stark, A., Roy, S. & Kellis, M. Reliable prediction of regulator targets using 12 *Drosophila* genomes. *Genome Res.* **17**, 1919–1931 (2007).
41. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature Biotechnol.* **28**, 817–825 (2010).
42. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43–49 (2011).
43. Pillas, D. *et al.* Genome-wide association study reveals multiple loci associated with primary tooth development during infancy. *PLoS Genet.* **6**, e1000856 (2010).
44. Lowe, C. B., Bejerano, G. & Haussler, D. Thousands of human mobile element fragments undergo strong purifying selection near developmental genes. *Proc. Natl Acad. Sci. USA* **104**, 8005–8010 (2007).
45. Prabhakar, S. *et al.* Human-specific gain of function in a developmental enhancer. *Science* **321**, 1346–1350 (2008).
46. Pollard, K. S. *et al.* Forces shaping the fastest evolving regions in the human genome. *PLoS Genet.* **2**, e168 (2006).
47. Schmidt, D. *et al.* Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328**, 1036–1040 (2010).
48. Mikkelsen, T. S. *et al.* Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* **447**, 167–177 (2007).
49. Genome 10K Community Of Scientists. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J. Hered.* **100**, 659–674 (2009).
50. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank O. Ryder, E. Fuchs, D. Haring, A. Walsh, D. Duffield, S. Wong, T. Alvarado, J. Boylan, S. Combes, P. deJong, J. Allman, J. Patton, D. McMullen, D. Hafner, D. Miller, T. Kunz, G. Hewitt, J. Searle, H. Künzle and D. Williams for providing organismal material. We thank L. Gaffney for help with figures. This work was supported by the National Human Genome Research Institute (NHGRI), including grant U54 HG003273 (R.A.G.), National Institute for General Medicine (NIGMS) grant no.

GM82901 (Pollard laboratory) and the European Science Foundation (EURYI award to K.L.-T.), NSF National Science Foundation (NSF) postdoctoral fellowship award 0905968 (J.E.), National Science Foundation CAREER 0644282 and NIH R01 HG004037 and the Sloan Foundation (M.K.), and an Erwin Schrödinger Fellowship of the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung (S.W.), the Gates Cambridge Trust (G.J.), Novo Nordisk Foundation (B.J.P. and J.W.); a Statistics Network Fellowship, Department of Mathematical Sciences, University of Copenhagen (B.J.P.); the David and Lucile Packard Foundation (A.S.); the Danish Council for Independent Research Medical Sciences (J.S.P.); The Lundbeck Foundation (J.S.P.).

**Author Contributions** K.L.-T., E.S.L. and M.K. led the project and oversaw the analysis. K.L.-T., M.C., J.Ch., E.H.M., E.D.G. and E.S.L. planned the project. K.L.-T., F.D.P., M.L., E.S.L., K.C.W., C.L.K., D.M.M., R.A.G., W.C.W., E.R.M., G.M.W. and R.K.W. oversaw or significantly contributed to data generation. S.G. assembled the 2 × genomes. Major contributions to analysis were made by M.Ga., O.Z. and M.C. to evaluate measures and patterns of evolutionary selection, M.F.L. to evaluate protein-coding potential and translational readthrough, B.J.P., S.W. to analyse RNA structures and families, P.K. on regulatory motifs and motif instances, J.E. on chromatin states, G.J. on codon-specific positive selection, E.M. on promoter motif, L.D.W. on GWAS overlap with conserved elements, C.B.L. on exaptation and A.K.H., K.S.P. on HARs and PARs. J.A., K.B., H.C., J.Cu., S.F., P.F., M.Gu., M.J.H., D.B.J., I.J., W.J.K., D.K., A.L.M., T.M., I.M., B.J.R., M.D.R., J.R., A.St., A.J.V., J.W., X.X., M.C.Z., E.B., E.H.M., J.H., D.H., A.Si., N.G. and J.S.P. performed or oversaw various analyses. K.L.-T., E.S.L. and M.K. wrote the paper with input from the other authors.

**Author Information** A complete set of data files can be downloaded from the Broad website (<https://www.broadinstitute.org/scientific-community/science/projects/mammals-models/29-mammals-project-supplementary-info>) or viewed using the UCSC Genome Browser (<http://genomewiki.cse.ucsc.edu/index.php/29mammals>) or the Broad Institute Integrative Genome Viewer (<http://www.broadinstitute.org/igv/projects/29mammals>). NCBI accession numbers for all newly sequenced genomes can be found in Supplementary Table 1. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at [www.nature.com/nature](http://www.nature.com/nature). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to K.L.-T. ([kersli@broadinstitute.org](mailto:kersli@broadinstitute.org)), E.S.L. ([lander@broadinstitute.org](mailto:lander@broadinstitute.org)) or M.K. ([manoli@mit.edu](mailto:manoli@mit.edu)).

#### Broad Institute Sequencing Platform and Whole Genome Assembly Team

Jen Baldwin<sup>1</sup>, Toby Bloom<sup>1</sup>, Chee Whye Chin<sup>1</sup>, Dave Heiman<sup>1</sup>, Robert Nicol<sup>1</sup>, Chad Nusbaum<sup>1</sup>, Sarah Young<sup>1</sup> & Jane Wilkinson<sup>1</sup>

#### Baylor College of Medicine Human Genome Sequencing Center Sequencing Team

Andrew Cree<sup>2</sup>, Huyen H. Dinh<sup>2</sup>, Gerald Fowler<sup>2</sup>, Shalili Jhangiani<sup>2</sup>, Vandita Joshi<sup>2</sup>, Sandra Lee<sup>2</sup>, Lora R. Lewis<sup>2</sup>, Lynne V. Nazareth<sup>2</sup>, Geoffrey Okwuonu<sup>2</sup> & Jireh Santibanez<sup>2</sup>

#### Genome Institute at Washington University

Kim Delehaunty<sup>3</sup>, David Dooling<sup>3</sup>, Catrina Fronik<sup>3</sup>, Lucinda Fulton<sup>3</sup>, Bob Fulton<sup>3</sup>, Tina Graves<sup>3</sup>, Patrick Minx<sup>3</sup> & Erica Sodergren<sup>3,4</sup>

<sup>1</sup>Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), 320 Charles Street, Cambridge, Massachusetts 02142, USA. <sup>2</sup>Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. <sup>3</sup>Genome Institute at Washington University, Washington University School of Medicine, 4444 Forest Park Blvd, Saint Louis, Missouri 63108, USA. <sup>4</sup>Research Institute of Molecular Pathology (IMP), A-1030 Vienna, Austria.

# Spatio-temporal transcriptome of the human brain

Hyo Jung Kang<sup>1\*</sup>, Yuka Imamura Kawasawa<sup>1\*</sup>, Feng Cheng<sup>1\*</sup>, Ying Zhu<sup>1\*</sup>, Xuming Xu<sup>1\*</sup>, Mingfeng Li<sup>1\*</sup>, André M. M. Sousa<sup>1,2</sup>, Mihovil Pletikos<sup>1,3</sup>, Kyle A. Meyer<sup>1</sup>, Goran Sedmak<sup>1,3</sup>, Tobias Guennel<sup>4</sup>, Yurae Shin<sup>1</sup>, Matthew B. Johnson<sup>1</sup>, Željka Krsnik<sup>1</sup>, Simone Mayer<sup>1,5</sup>, Sofia Fertuzinhos<sup>1</sup>, Sheila Umlauf<sup>6</sup>, Steven N. Lisgo<sup>7</sup>, Alexander Vortmeyer<sup>8</sup>, Daniel R. Weinberger<sup>9</sup>, Shrikant Mane<sup>6</sup>, Thomas M. Hyde<sup>9,10</sup>, Anita Huttner<sup>8</sup>, Mark Reimers<sup>4</sup>, Joel E. Kleinman<sup>9</sup> & Nenad Šestan<sup>1</sup>

**Brain development and function depend on the precise regulation of gene expression. However, our understanding of the complexity and dynamics of the transcriptome of the human brain is incomplete. Here we report the generation and analysis of exon-level transcriptome and associated genotyping data, representing males and females of different ethnicities, from multiple brain regions and neocortical areas of developing and adult post-mortem human brains. We found that 86 per cent of the genes analysed were expressed, and that 90 per cent of these were differentially regulated at the whole-transcript or exon level across brain regions and/or time. The majority of these spatio-temporal differences were detected before birth, with subsequent increases in the similarity among regional transcriptomes. The transcriptome is organized into distinct co-expression networks, and shows sex-biased gene expression and exon usage. We also profiled trajectories of genes associated with neurobiological categories and diseases, and identified associations between single nucleotide polymorphisms and gene expression. This study provides a comprehensive data set on the human brain transcriptome and insights into the transcriptional foundations of human neurodevelopment.**

Human neurodevelopment is a complex and precisely regulated process that occurs over a protracted period of time<sup>1–3</sup>. Human-specific features of this process are likely to be important factors in the evolution of human specializations<sup>2–5</sup>. However, in addition to giving us remarkable cognitive and motor abilities, the formation of molecularly distinct and intricate neural circuits may have also increased our susceptibility to certain psychiatric and neurological disorders<sup>4–9</sup>. Furthermore, sex differences are important in brain development and function, and are a risk factor for conditions such as autism spectrum disorders (ASDs) and depression<sup>9–13</sup>. Research and progress in all these areas could be enhanced by a comprehensive analysis of the spatio-temporal dynamics of gene expression and transcript variants in the human brain.

Previous transcriptome studies of the developing human brain have used relatively small numbers of samples and predominantly focused on only a few regions or developmental time points<sup>14–18</sup>. In this Article, we explore the transcriptomes of 16 regions comprising the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus and 11 areas of the neocortex. The data set was generated from 1,340 tissue samples collected from 57 developing and adult post-mortem brains of clinically unremarkable donors representing males and females of multiple ethnicities.

## Study design, data generation and quality control

To investigate the spatio-temporal dynamics of the human brain transcriptome, we created a 15-period system spanning the periods from embryonic development to late adulthood (Table 1 and Supplementary Information, section 2.1). We sampled transient prenatal

structures and immature and mature forms of 16 brain regions, including 11 neocortex (NCX) areas, from multiple specimens per period (Table 2; Supplementary Information, section 2.2; Supplementary Figs 1–3; and Supplementary Table 1). The 11 NCX areas are collectively referred to hereafter as the region NCX. We also genotyped donor DNA using an Illumina 2.5-million single nucleotide polymorphism (SNP) chip (Supplementary Fig. 4 and Supplementary Table 2). Only brains from clinically unremarkable donors with no signs of large-scale genomic abnormalities were included in the study ( $N = 57$ , including 39 with both hemispheres; age, 5.7 weeks post-conception

**Table 1 | Periods of human development and adulthood as defined in this study**

Period	Description	Age
1	Embryonic	4 PCW ≤ Age < 8 PCW
2	Early fetal	8 PCW ≤ Age < 10 PCW
3	Early fetal	10 PCW ≤ Age < 13 PCW
4	Early mid-fetal	13 PCW ≤ Age < 16 PCW
5	Early mid-fetal	16 PCW ≤ Age < 19 PCW
6	Late mid-fetal	19 PCW ≤ Age < 24 PCW
7	Late fetal	24 PCW ≤ Age < 38 PCW
8	Neonatal and early infancy	0 M (birth) ≤ Age < 6 M
9	Late infancy	6 M ≤ Age < 12 M
10	Early childhood	1 Y ≤ Age < 6 Y
11	Middle and late childhood	6 Y ≤ Age < 12 Y
12	Adolescence	12 Y ≤ Age < 20 Y
13	Young adulthood	20 Y ≤ Age < 40 Y
14	Middle adulthood	40 Y ≤ Age < 60 Y
15	Late adulthood	60 Y ≤ Age

M, postnatal months; PCW, post-conceptional weeks; Y, postnatal years.

<sup>1</sup>Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA. <sup>2</sup>Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, 4099-003 Porto, Portugal. <sup>3</sup>Graduate Program in Neuroscience, Croatian Institute for Brain Research, University of Zagreb School of Medicine, 10000 Zagreb, Croatia. <sup>4</sup>Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia 23298, USA. <sup>5</sup>MSc/PhD Molecular Biology Program, International Max Planck Research School for Molecular Biology, 37077 Göttingen, Germany. <sup>6</sup>Yale Center for Genome Analysis, Yale University School of Medicine, New Haven, Connecticut 06510, USA. <sup>7</sup>Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle upon Tyne NE1 3BZ, UK. <sup>8</sup>Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06510, USA. <sup>9</sup>Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>10</sup>The Lieber Institute for Brain Development, Johns Hopkins University Medical Campus, Baltimore, Maryland 21205, USA.

\*These authors contributed equally to this work.

**Table 2 | Ontology and nomenclature of analysed brain regions and NCX areas**

Periods 1 and 2	Periods 3–15
FC, frontal cerebral wall	OFC, orbital prefrontal cortex DFC, dorsolateral prefrontal cortex VFC, ventrolateral prefrontal cortex MFC, medial prefrontal cortex M1C, primary motor (M1) cortex
PC, parietal cerebral wall	S1C, primary somatosensory (S1) cortex IPC, posterior inferior parietal cortex
TC, temporal cerebral wall	A1C, primary auditory (A1) cortex STC, superior temporal cortex ITC, inferior temporal cortex
OC, occipital cerebral wall	V1C, primary visual (V1) cortex
HIP, hippocampal anlage	HIP, hippocampus
—	AMY, amygdala
VF, ventral forebrain MGE, medial ganglionic eminence LGE, lateral ganglionic eminence CGE, caudal ganglionic eminence	STR, striatum
DIE, diencephalon DTH, dorsal thalamus	MD, mediodorsal nucleus of the thalamus —
URL, upper (rostral) rhombic lip	CBC, cerebellar cortex

(PCW) to 82 years; sex, 31 males and 26 females; post-mortem interval,  $12.11 \pm 8.63$  (mean  $\pm$  s.d.) hours; pH,  $6.45 \pm 0.34$  (mean  $\pm$  s.d.).

Transcriptome profiling was performed using total RNA extracted from a total of 1,340 dissected tissue samples (RNA integrity number,  $8.83 \pm 0.93$  (mean  $\pm$  s.d.); Supplementary Tables 3 and 4). We used the Affymetrix GeneChip Human Exon 1.0 ST Array platform, which features comprehensive coverage of the human genome, with 1.4 million probe sets that assay expression across the entire transcript or individual exon, thereby providing redundancy and increased confidence in estimates of gene-level differential expression (DEX, differentially expressed) and differential exon usage (DEU). Descriptions of tissue sampling and quality control measures implemented throughout transcriptome data generation steps are provided in Supplementary Information, sections 2–5, and Supplementary Figs 5–8.

## Global transcriptome dynamics

### Spatio-temporal gene expression

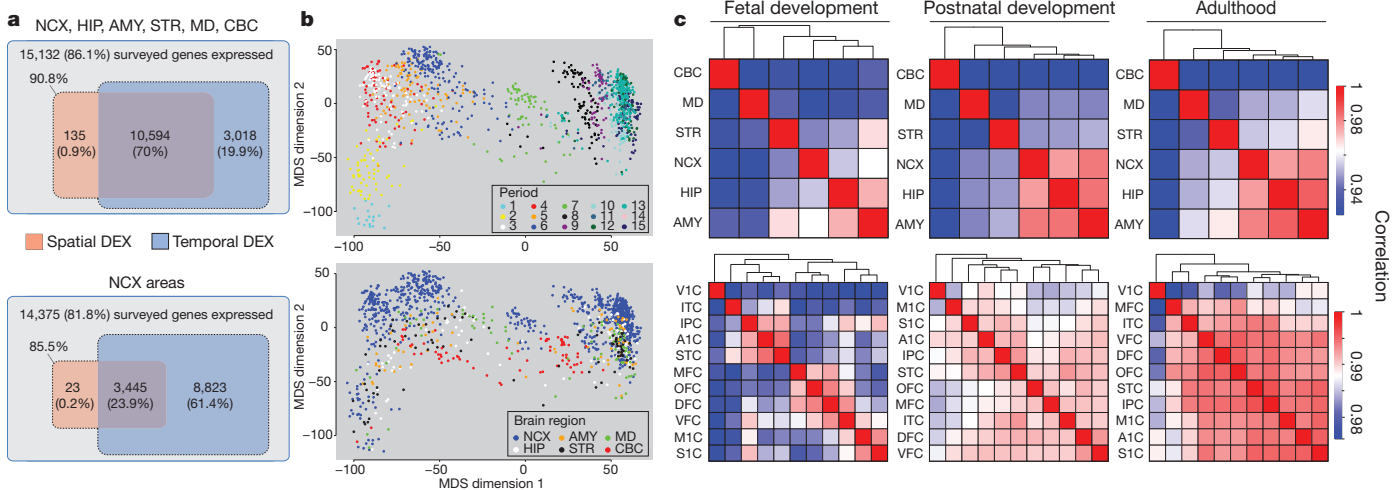
After quality control assessments and quantile normalization, we summarized core and unique probe sets, representing 17,565 mainly

protein-coding genes, into gene-level information. Using stringent criteria ( $\log_2$ -transformed signal intensity of  $\geq 6$  in at least one sample and mean detection-above-background  $P$  value of  $<0.01$  in at least one region of at least one period) to define an ‘expressed’ gene, we found that 15,132 (86.1%) of 17,565 genes surveyed were expressed in at least one brain region during at least one period, and that 14,375 (81.8%) were expressed in at least one NCX area (Fig. 1a, Supplementary Information 6.1 and Supplementary Fig. 9). To investigate the contributions of different factors to the global transcriptome dynamics, we applied multidimensional scaling and principal-component analysis, which revealed that region and age (that is, spatio-temporal dynamics) contribute more to the global differences in gene expression than do other tested variables: sex, ethnicity and inter-individual variation (Fig. 1b; Supplementary Information, sections 6.2 and 6.3; and Supplementary Figs 10 and 11).

To identify genes that were spatially or temporally regulated, we used a conservative threshold (false-discovery-rate  $Q$  value of  $<0.01$  and  $\geq 2$ -fold  $\log_2$ -transformed signal intensity difference), included post-mortem interval and RNA integrity number as technical covariates within all of our analysis-of-variance models of differential expression, considered the influence of dissection variation and applied a fivefold jackknife procedure (Supplementary Information, section 6.4, and Supplementary Figs 12 and 13). We found that 70.9% of expressed genes were spatially DEX between any two regions within at least one period, and that 24.1% were spatially DEX between any two NCX areas (Fig. 1a). By contrast, 89.9% of expressed genes were temporally DEX between any two periods across regions, and 85.3% were temporally DEX between any two periods across NCX areas. Moreover, 70.0% and 23.9% of expressed genes were both spatially and temporally DEX within brain regions and within NCX areas, respectively. The bulk of spatio-temporal regulation occurred during prenatal development. For instance, 57.7% of NCX-expressed genes were temporally DEX across fetal development (periods 3–7), whereas 9.1% were during postnatal development (periods 8–12) and 0.7% were during adulthood (periods 13–15). Together, these data indicate that the majority of brain-expressed protein-coding genes are temporally and, to a lesser extent, spatially regulated, and that this regulation occurs predominantly during prenatal development.

### Transcriptional architecture of the human brain

To assess transcriptional relatedness between brain regions/NCX areas, we calculated correlation matrices of pairwise comparisons



**Figure 1 | Global spatio-temporal dynamics of gene expression.** **a**, Venn diagrams representing the total number of genes considered to be expressed and the number of spatially and temporally DEX genes for brain regions (top) and NCX areas (bottom). **b**, Multidimensional scaling (MDS) plot showing transcriptional similarity, coloured by period (top) and region (bottom). Non-

metric; stress = 18.9%. Euclidean distance of  $\log_2$ -transformed signal intensity was used to measure pairwise similarity. **c**, Heat map matrix of pairwise Spearman correlations between brain regions (top) and between NCX areas (bottom) during fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15).

(Fig. 1c) and performed unsupervised hierarchical clustering across periods 3–15, an interval during which all analysed regions/areas can be consistently followed across time. Among regions, this analysis showed distinct and developmentally regulated clustering of NCX (combination of 11 areas), HIP and AMY, with CBC having the most distinctive transcriptional profile. At the level of NCX areas, clustering formed the following groups during fetal periods: OFC, DFC and MFC; VFC and primary somatomotor cortex (S1C and M1C); and parietal-temporal perisylvian areas (IPC, A1C and STC). VIC had the most distinctive transcriptional profile of NCX areas throughout development and adulthood. The increased correlations between NCX, HIP, AMY and the majority of non-VIC NCX areas with age indicate that transcriptional differences are particularly pronounced during development.

Consistent with the clustering observed, CBC showed the greatest number of region-restricted or region-enriched DEX genes, with 516 (4.8%) of 10,729 genes spatially DEX (Supplementary Information, section 6.4, and Supplementary Table 5). By contrast, the numbers of genes highly enriched in the other regions were lower: NCX, 46 (0.43%); HIP, 48 (0.45%); AMY, 4 (0.04%); STR, 137 (1.28%); MD, 216 (2.01%). The majority of these spatially enriched genes were also temporally regulated, and some, such as those in Supplementary Figs 14 and 15 (NCX: *FLJ32063*, *KCNS1*; HIP: *CDC20B*, *METTL7B*; AMY: *TFAP2D*, *UTS2D*; STR: *C10orf11*, *PTPN7*; MD: *CEACAM21*, *SLC24A5*; CBC: *ESRRB*, *ZP2*), were transiently enriched during a narrow time window. These clustering and region-enrichment results reveal that regional transcriptomes are developmentally regulated and reflect anatomical differences.

### Spatio-temporal differential exon usage

Alternative exon usage is an important mechanism for generating transcript diversity<sup>19,20</sup>. Using a splicing analysis of variance and a splicing index algorithm with conservative criteria ( $Q < 0.01$  with a minimum twofold splice index difference between at least two regions/areas or periods; Supplementary Information, section 6.5), we found that 13,647 (90.2%) of 15,132 expressed genes showed DEU across sampled regions (0.1%), periods (19.5%) or both (70.6%). Of 14,375 NCX genes, 88.7% showed DEU across sampled areas ( $<0.01\%$ ), periods (59.8%) or both (28.9%). The regulation of DEU also varied in time, with the majority of expressed genes (83.0%) showing temporal DEU across fetal development, whereas only 0.9% and 1.4% were temporally regulated across postnatal development and adulthood, respectively.

Focusing on *ANKRD32*, a gene we have previously shown to express an alternative variant in the late mid-fetal frontal cortex<sup>16</sup>, we

confirmed and extended our findings on DEU by showing that whereas the longer isoform (*ANKRD32a*) was equally expressed across fetal NCX areas, the shorter isoform (*ANKRD32b*), comprising the last three exons, exhibited dynamic areal patterns. *ANKRD32b* was transiently expressed in a gradient along the anterior–posterior axis of the mid-fetal frontal cortex, with the highest expression in OFC and the lowest in M1C. Before this, *ANKRD32b* was most highly enriched in the ITC and, to a lesser extent, the STC. These spatio-temporal patterns disappeared after birth, when only *ANKRD32a* was expressed, and were not observed in the mouse NCX of equivalent ages (Supplementary Fig. 16 and Supplementary Table 6). These findings illustrate the complexity of DEU in the human brain and demonstrate how specific alternative transcripts can be spatially restricted during a narrow developmental window and with interspecies differences.

### Sex differences in the transcriptome

#### Sex-biased gene expression

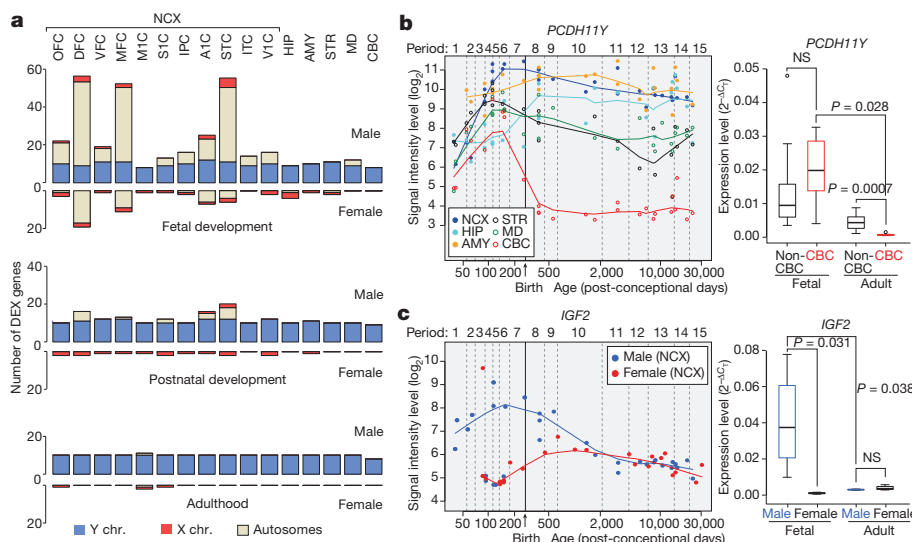
Previous studies have identified sexually dimorphic gene expression in the developing and adult human brain<sup>11–13</sup>. Analysis of our data set using a sliding-window algorithm and  $t$ -test model ( $Q < 0.01$  with  $>2$ -fold difference in  $\log_2$ -transformed signal intensity; Supplementary Information, section 6.6) identified 159 genes, including a number of previously reported and newly uncovered genes with male or female bias in expression located on the Y (13 genes), X (9 genes) and autosomal (137 genes) chromosomes. A large fraction (76.7%) had male-biased expression (Fig. 2a and Supplementary Table 7). Notable spatial differences were observed, and more genes had sex-biased expression during prenatal development than during postnatal life, with the adult brain characterized by having the fewest.

Consistent with previous findings<sup>12,13</sup>, we found that the largest differences were attributable to Y-chromosome genes, especially *PCDH11Y*, *RPS4Y1*, *USP9Y*, *DDX3Y*, *NLGN4Y*, *UTY*, *EIF1AY* and *ZFY*, which showed constant expression across regions and periods, with the exception of *PCDH11Y* downregulation in the postnatal CBC (Fig. 2b). Notably, the functional homologues of these genes on the X chromosome, barring *ZFX* during fetal development (*PCDH11X*, *RPS4X*, *USP9X*, *DDX3X*, *NLGN4X*, *UTX* and *EIF1AX*), were not upregulated in a compensatory manner in female brains (Supplementary Fig. 17).

We also found other X-linked and autosomal genes with sex-biased expression and distinct spatio-temporal patterns, including functionally uncharacterized transcripts (*LOC554203*, *C3orf62*, *FLJ35409* (also known as *MIR137HG*) and *DKFZP586I1420*), *S100A10* (which has been linked to depression<sup>21</sup>) and *IGF2* (an imprinted autosomal gene previously implicated in embryonic growth and cognitive function<sup>22,23</sup>), that showed population-level male-biased expression (Fig. 2c).

**Figure 2 | Sex-biased gene expression.**

**a**, Number of sex-biased DEX genes in brain regions/NCX areas during fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15). **b**, *PCDH11Y* exon array signal intensity (left) and validation by quantitative PCR with reverse transcription (qRT-PCR; right) ( $N = 5$  male brains per period). **c**, *IGF2* exon array signal intensity (left) and qRT-PCR (right) validation in NCX ( $N = 4$  per sex and period).  $P$  values were calculated by unpaired  $t$ -test. Whiskers indicate fifth and ninety-fifth percentiles, respectively. NS, not significant.



## Sex-biased exon usage

We next explored sex-biased DEU using a sliding-window algorithm with a splicing *t*-test model ( $Q < 0.01$  and splicing index  $> 2$ ; Supplementary Information, section 6.6). We identified 155 genes (145 autosomal) that showed sex-biased expression of probe sets encoding one or a subset of exons (Supplementary Table 8) in one or multiple regions/NCX areas. These included several members of the collagen family of genes (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A2* and *COL6A3*), *C3*, *KCNH2* (a gene associated with schizophrenia<sup>24</sup>), *NOTCH3* (a gene mutated in a common form of hereditary stroke disorder<sup>25</sup>), *ELN* (a gene located within the Williams syndrome critical region<sup>26</sup>) and *NLGN4X* (an X-chromosome gene implicated in synapse function and associated with ASD and moderate X-linked intellectual disability<sup>10,27</sup>). Although comparably expressed in males and females at the population and gene levels (Supplementary Fig. 17), *NLGN4X* had a significant male bias in expression of exon 7 and, to a lesser extent, exons 1, 5 and 6 in a developmentally regulated manner (Fig. 3). Together, these findings show that developmentally and spatially regulated differences in gene- and exon-level expression exist between male and female brains.

## Gene co-expression networks

To extract additional biological information embedded in the multi-dimensional transcriptome data set, we performed weighted gene co-expression network analysis<sup>28</sup>, which allowed us to identify modules of co-expressed genes. We identified 29 modules associated with distinct spatio-temporal expression patterns and biological processes (Fig. 4a; Supplementary Information, section 6.7; Supplementary Tables 9–11; and Supplementary Figs 18–20). Among modules corresponding to specific spatio-temporal patterns, M8 consisted of 24 genes with a common developmental trend that showed the highest expression levels in early fetal NCX and HIP (period 3), followed by a progressive decline in expression levels with age until infancy (period 9) (Fig. 4b). By contrast, M15 contained 310 genes showing changes in the opposite direction (relative to those in M8) in the NCX, HIP, AMY and STR (Fig. 4c). Gene ontology enrichment analysis showed that genes in M8 were enriched for gene ontology categories related to neuronal differentiation (Bonferroni-adjusted  $P = 7.7 \times 10^{-3}$ ) and

transcription factors ( $P = 5.2 \times 10^{-3}$ ) (Supplementary Information, section 6.8, and Supplementary Table 9). Conversely, M15 gene ontology categories included ionic channels ( $P = 8.0 \times 10^{-8}$ ) and neuroactive ligand–receptor interaction ( $P = 4.0 \times 10^{-14}$ ).

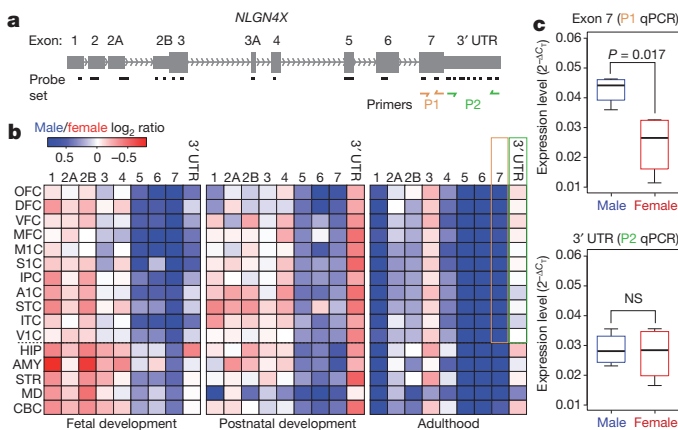
Genes with the highest degree of connectivity within a module are termed hub genes and are expected to be functionally important within the module. M8 hub genes included transcription factors *TBR1*, *FEZF2*, *FOXG1*, *SATB2*, *NEUROD6* and *EMX1* (Fig. 4b), which have been functionally implicated in the development of NCX and HIP projection neurons<sup>29–38</sup>. Furthermore, *FOXG1* variants have been linked to Rett syndrome and intellectual disability<sup>34</sup>. Sequence variants in M15 hub genes (Fig. 4c) have been linked to major depression<sup>39</sup> (*GDA*) and to schizophrenia and affective disorders<sup>6,40</sup> (*NRGN* and *RGS4*).

We also identified two large-scale, temporally regulated modules (M20 and M2) with opposite developmental trajectories of genes co-expressed across regions: expression in M20 gradually decreased with age and expression in M2 gradually increased (Supplementary Figs 21 and 22). M20 was enriched for gene ontology categories related to zinc-finger proteins ( $P = 7.3 \times 10^{-48}$ ) and transcription factors ( $P = 4.8 \times 10^{-50}$ ), including many ZNF and SOX family members. M2 was enriched for gene ontology categories related to membrane proteins ( $P = 1.8 \times 10^{-21}$ ), calcium signalling ( $P = 8.1 \times 10^{-10}$ ), synaptic transmission ( $P = 1.6 \times 10^{-6}$ ) and neuroactive ligand–receptor interaction ( $P = 4.1 \times 10^{-4}$ ), reflecting processes important in postnatal brain maturation. Their hub genes encoded transcriptional factors, modulators of chromatin state and signal transduction proteins, all of which are likely to be involved in driving the co-expression networks. Drastic expression shifts in M20 and M2 in the opposite direction just before birth indicate that this period is associated with global transcriptional changes that probably reflect environmental influences on brain development and intrinsic changes in cellular composition and functional processes.

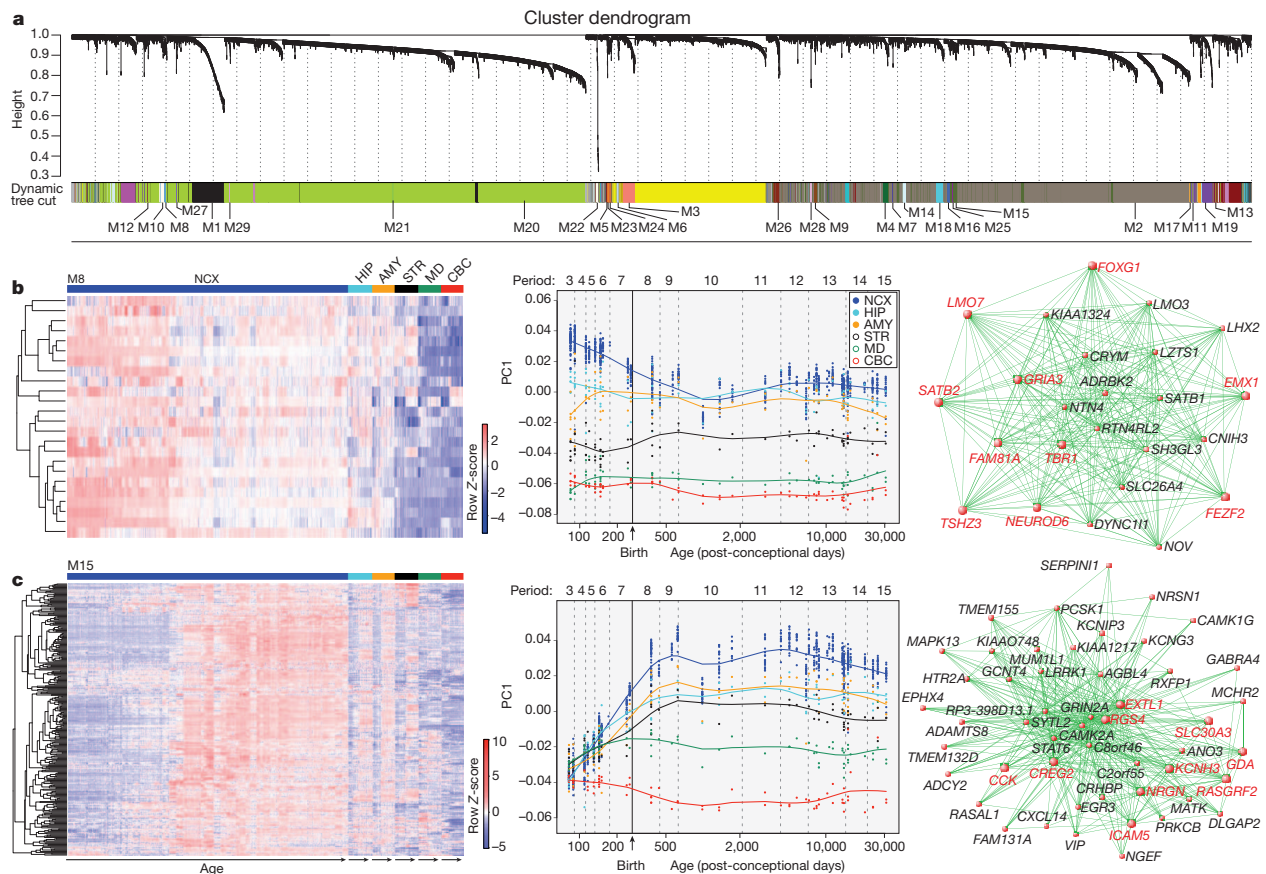
## Expression trajectories of neurodevelopment

One important use for the generated data set is to gain insight into normal and abnormal human neurodevelopment by analysing trajectories of individual genes or groups of genes associated with a particular neurobiological category or disease. To test this strategy, we compared our expression data for *DCX* (a gene expressed in neuronal progenitor cells and immature migrating neurons), as well as for genes associated with dendrite (*MAP1A*, *MAPT*, *CAMK2A*) and synapse (*SYN*, *SYPL1*, *SYPL2*, *SYN1*) development, with independently generated, non-transcriptome human data sets. The *DCX* expression trajectory was remarkably reminiscent of the reported changes in the density of *DCX*-immunopositive cells in the postnatal human HIP<sup>36,40</sup> ( $r = 0.946$ , Pearson correlation; Fig. 5a). In our transcriptome data set, *DCX* expression increased until early mid-fetal development (period 5) and then gradually declined with age until early childhood (period 10). Likewise, expression trajectories of dendrite and synapse development gene groups closely paralleled the growth of basal dendrites of DFC pyramidal neurons<sup>41</sup> ( $r = 0.810$  for layer 3 and  $r = 0.700$  for layer 5; Fig. 5b) and DFC synaptogenesis<sup>42</sup> ( $r = 0.940$ ; Fig. 5c), respectively. Steep increases in both processes occurred between the late mid-fetal period and late infancy, indicating that a considerable portion of these two processes occurs before birth and reaches a plateau around late infancy.

After demonstrating the accuracy and viability of using the data set to profile human neurodevelopment, we manually curated lists of genes associated with over 80 categories, including various neurodevelopmental processes, neural cell types and neurotransmitter systems (Supplementary Information 6.9 and Supplementary Table 12). Notable trajectories and differences in their onset times, rates of increase and decrease, and shapes were observed within and between brain regions for categories including major neurodevelopmental processes (neural cell proliferation and migration, dendrite and synapse development, and myelination; Fig. 5d), cortical GABAergic inhibitory interneurons (*CALB1*, *CALB2*, *NOS1*, *PVALB* and *VIP*) and



**Figure 3 | Sex-biased differential exon usage.** **a**, Gene structure and probe set composition of *NLGN4X*. Yellow and green arrows depict primers used for qRT-PCR validation. **b**, Heat map of the  $\log_2$  male/female signal intensity ratio of each exon for fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15). Differences in expression of exon 7 (yellow frame) and the 3' untranslated region (UTR; green frame) in adult NCX are highlighted. Note that exons 2 and 3A did not meet our expression criteria and are not represented. **c**, qRT-PCR validation of expression of exon 7 and the 3' UTR in adult NCX ( $N = 4$  per sex).  $P$  values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively.



**Figure 4 | Global co-expression networks and gene modules.** **a**, Dendrogram from gene co-expression network analysis of samples from periods 3–15. Modules of co-expressed genes were assigned a colour and number (M1 to M29). **b**, Left: heat map of genes in M8 showing the spatio-temporal expression pattern after hierarchical clustering. The expression values for each gene are arranged in the heat map, ordered first by brain region, then by age and last by

NCX area. Middle: spatio-temporal pattern of M8 summarized by the first principal component (PC1) for expression of genes in the module across age. Right: 24 M8 genes; the top ten hub genes are shown in red. **c**, Same analyses as in **b**, but for M15; the top 50 genes defined by the highest intramodular connectivity are shown in right panel. Results for other modules are available in Supplementary Information.

glutamate receptors (Supplementary Figs 23 and 24). Two expected patterns were observed in neurodevelopmental trajectories: changes in expression of cell proliferation genes preceded the increase in expression of *DCX*, and expression of each decreased during perinatal development whereas synapse development, dendrite development and myelination trajectories increased. Notably, the NCX trajectory for synapse development did not drastically decline during late childhood or adolescence (Fig. 5c, d) as previously reported for synapse density<sup>42</sup>. We also identified co-expression network modules and additional genes that are highly correlated with the categories (Supplementary Tables 10, 13 and 14). For example, M20 and M2 were strongly correlated with neuron migration ( $r = 0.894$ ) and myelination ( $r = 0.972$ ), respectively.

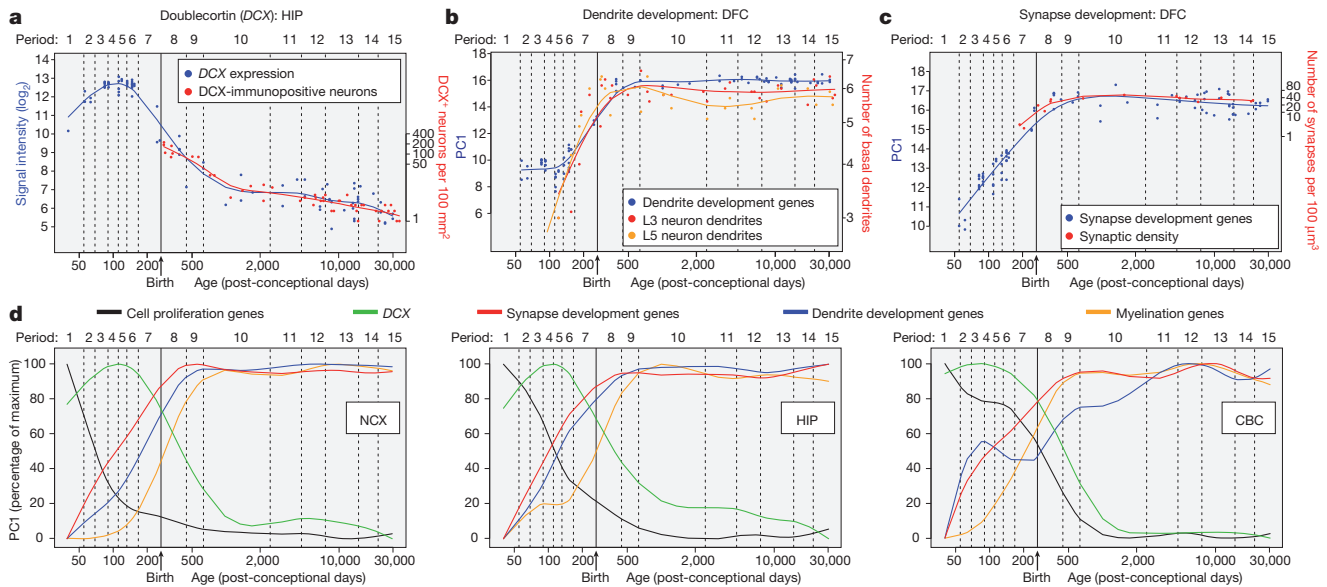
In addition, our data set enabled us to generate expression trajectories of genes commonly associated with ASD and schizophrenia. We investigated a number of genes previously linked to these disorders (Supplementary Information, section 6.10) and observed distinct and dynamic expression patterns, especially among NCX areas (Supplementary Fig. 25 shows examples for *CNTNAP2*, *MET*, *NLGN4X* and *NRGN*). To gain insight into potential biological functions of ASD- and schizophrenia-associated genes in human neurodevelopment, we identified other genes with significantly correlated spatio-temporal expression profiles and performed gene ontology enrichment analysis (Supplementary Tables 15 and 16). These findings reveal associated spatio-temporal differences in these expression trajectories and provide additional co-expressed genes that can be interrogated for their role in the respective processes or disorders.

## Expression quantitative trait loci

Previous studies have identified expression quantitative trait loci (eQTLs) in the adult human brain, primarily in the cerebral cortex<sup>43–47</sup>. Our multiregional developmental data set enabled us to search for association between SNP genotypes and spatio-temporal gene expression. We tested only for *cis*-eQTLs, restricting the search to SNPs within 10 kilobases of either a transcription start site or a transcription end site, as opposed to *trans*-eQTLs, which would require much larger sample sizes.

Implementing a conservative strategy (gene-wide Bonferroni correction followed by genome-wide  $Q < 0.1$ ; Supplementary Information, section 9), we identified 39 NCX, eight HIP, four AMY, two STR, six MD and five CBC genes (Supplementary Table 17) with evidence of *cis*-eQTL, including two previously reported genes<sup>45,47</sup> (*ITGB3BP* and *ANKRD27*). Consistent with previous studies<sup>48</sup>, associated SNPs were enriched near transcription start and termination sites (Fig. 6a, b).

An example of a significant association in NCX, MD and CBC is that between SNP rs10785190 and *GLIPRIL2*, a member of the glioma pathogenesis-related 1 family of genes<sup>49</sup>. The expression differences were observed at the level of the whole transcript and exons 1 and 2, the only exons we observed to be expressed at appreciable levels in the NCX (Fig. 6c, d). The NCX probably had more *cis*-eQTLs than other regions owing to its smaller variation in gene expression resulting from the averaged expression of 11 areas. Many eQTLs identified as significant in NCX also have similar associations in other regions, although they were not statistically significant after the conservative genome-wide correction (Supplementary Table 17). Thus, we have



**Figure 5 | Trajectories of genes associated with neurodevelopmental processes.** **a**, Comparison between *DCX* expression in HIP and the density of *DCX*-immunopositive cells in the human dentate gyrus<sup>36</sup>. **b**, Comparison between transcriptome-based dendrite development trajectory in DFC and Golgi-method-based growth of basal dendrites of layer 3 (L3) and 5 (L5) pyramidal neurons in the human DFC<sup>41</sup>. **c**, Comparison between transcriptome-based synapse development trajectory in DFC and density of

DFC synapses calculated using electron microscopy<sup>42</sup>. For **b** and **c**, PC1 for gene expression was plotted against age to represent the developmental trajectory of genes associated with dendrite (**b**) or synapse (**c**) development. Independent data sets were centred, scaled and plotted on a logarithmic scale. **d**, PC1 value for the indicated sets of genes (expressed as percentage of maximum) plotted against age to represent general trends and regional differences in several neurodevelopmental processes in NCX, HIP and CBC.

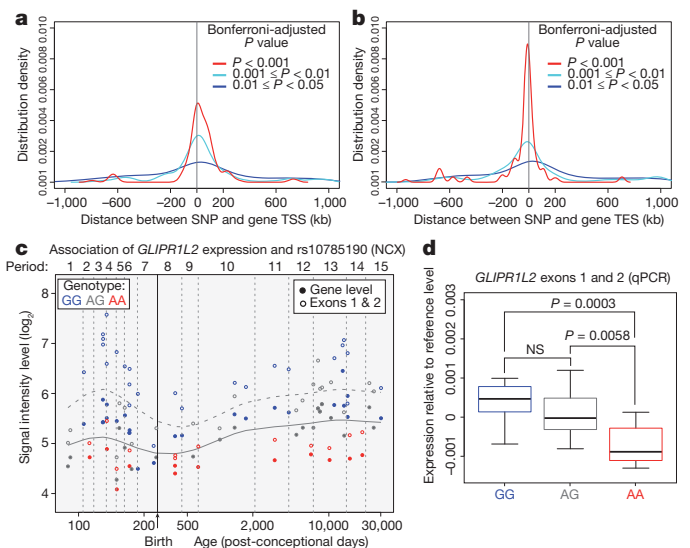
identified polymorphic regulators of transcription in different regions across development, potentially providing insights into inter-individual differences and genetic control of the brain transcriptome.

## Discussion

Our analysis reveals several features of the human brain transcriptome, and increases our knowledge of the transcriptional events in human neurodevelopment. We show that gene expression and exon usage have complex and dynamically regulated patterns, some of which may not be evident in the transcriptomes of commonly studied model organisms. Moreover, these patterns differ more prominently across time and space than they do between sexes, ethnicities or individuals, despite their underlying genetic differences. Transcriptome differences between males and females also included several disease-related genes, offering possible mechanisms underlying the sex differences in the incidence, prevalence and severity of some brain disorders. We also found that some of the inter-individual variations in the regional and developmental transcriptomes were associated with specific SNP genotypes, which may have altered expression-regulating elements. Thus, the present data set (available at <http://www.humanbraintranscriptome.org>), along with an accompanying study<sup>50</sup>, provides a basis for a variety of further investigations and comparisons with other transcriptome-related data sets of both healthy and diseased states.

Although our study has uncovered many intricacies in gene expression and exon usage in the human brain, there are potential limitations of our study that warrant discussion. Foremost, we used stringent criteria to minimize false positives and faithfully characterize general transcriptional patterns, rather than to capture all the changes that may occur. Also, we analysed dissected tissue that contained multiple cell types, thus diluting the transcriptional contribution and dynamic range of expression of any one specific cell type. Current limitations prevent us from using cell-type-specific approaches in systematically analysing the spatio-temporal transcriptome. Furthermore, the number of brains and regions analysed so far is not sufficient to investigate the full magnitude of transcriptional changes or the full range of eQTLs. Application of sequencing technology will allow more in-depth analyses of the transcriptome, and aid in discovery of novel or

low-expressing transcripts. Finally, although specific patterns of expression are often linked to specialized biological processes, it is important to remember that the relationship between messenger RNA and protein levels is not always linear nor translated into apparent phenotypic differences. As these concerns are addressed in future



**Figure 6 | Association between SNPs and gene expression.** **a**, SNP distribution around transcription start sites (TSS; **a**) and transcription end sites (TES; **b**) of the associated genes, based on several Wald test *P*-value cut-offs after gene-wide Bonferroni correction. **c**, *GLIPR1L2* expression association with rs10785190 genotype, a SNP located in exon 1. The solid and dashed curves, calculated from locally weighted scatter-plot smoothing (LOWESS), show the developmental trends of gene expression and exon-1 and exon-2 expression, respectively. **d**, qRT-PCR validation of exon-1 and exon-2 expression in NCX for each genotype (*N* = 14 GG, 14 AG, 8 AA), plotted relative to the LOWESS curve in **c** to facilitate comparison across developmental periods. *P* values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively.

with more samples and new data sets from human and non-human primate brains, it will be possible to uncover deeper insights into the transcriptional foundations of human brain development and evolution.

## METHODS SUMMARY

Supplementary Information, sections 3–9, provides a full description of tissue acquisition and processing, data generation, validation and analyses.

Received 11 December 2010; accepted 30 August 2011.

- Kostovic, I. & Judas, M. Prolonged coexistence of transient and permanent circuitry elements in the developing cerebral cortex of fetuses and preterm infants. *Dev. Med. Child Neurol.* **48**, 388–393 (2006).
- Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nature Rev. Neurosci.* **10**, 724–735 (2009).
- Rubenstein, J. L. Annual Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. *J. Child Psychol. Psychiatry* **52**, 339–355 (2011).
- Preuss, T., Cáceres, M., Oldham, M. & Geschwind, D. Human brain evolution: insights from microarrays. *Nature Rev. Genet.* **5**, 850–860 (2004).
- Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* **437**, 64–67 (2005).
- Lewis, D. A. & Levitt, P. Schizophrenia as a disorder of neurodevelopment. *Annu. Rev. Neurosci.* **25**, 409–432 (2002).
- Meyer-Lindenberg, A. & Weinberger, D. R. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Rev. Neurosci.* **7**, 818–827 (2006).
- Insel, T. Rethinking schizophrenia. *Nature* **468**, 187–193 (2010).
- State, M. The genetics of child psychiatric disorders: focus on autism and Tourette syndrome. *Neuron* **68**, 254–269 (2010).
- Jamain, S. *et al.* Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nature Genet.* **34**, 27–29 (2003).
- Vawter, M. *et al.* Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes. *Neuropsychopharmacology* **29**, 373–384 (2004).
- Weickert, C. *et al.* Transcriptome analysis of male-female differences in prefrontal cortical development. *Mol. Psychiatry* **14**, 558–561 (2009).
- Reinius, B. & Jazin, E. Prenatal sex differences in the human brain. *Mol. Psychiatry* **14**, 988–989 (2009).
- Abrahams, B. *et al.* Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc. Natl Acad. Sci. USA* **104**, 17849–17854 (2007).
- Sun, T. *et al.* Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science* **308**, 1794–1798 (2005).
- Johnson, M. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
- Somel, M. *et al.* MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* **20**, 1207–1218 (2010).
- Ip, B. *et al.* Investigating gradients of gene expression involved in early human cortical development. *J. Anat.* **217**, 300–311 (2010).
- Licatalosi, D. D. & Darnell, R. B. Splicing regulation in neurologic disease. *Neuron* **52**, 93–101 (2006).
- Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell* **126**, 37–47 (2006).
- Svenningsson, P. *et al.* Alterations in 5-HT<sub>1B</sub> receptor function by p11 in depression-like states. *Science* **311**, 77–80 (2006).
- Chen, D. Y. *et al.* A critical role for IGF-II in memory consolidation and enhancement. *Nature* **469**, 491–497 (2011).
- Lehtinen, M. K. *et al.* The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron* **69**, 893–905 (2011).
- Huffaker, S. J. *et al.* A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nature Med.* **15**, 509–518 (2009).
- Joutel, A. *et al.* Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383**, 707–710 (1996).
- Ewart, A. K. *et al.* Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. *Nature Genet.* **5**, 11–16 (1993).
- Südhof, T. C. Neuroligins and neuexins link synaptic function to cognitive disease. *Nature* **455**, 903–911 (2008).
- Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, article17 (2005).
- Hansen, D. V., Rubenstein, J. L. & Kriegstein, A. R. Deriving excitatory neurons of the neocortex from pluripotent stem cells. *Neuron* **70**, 645–660 (2011).
- Hevner, R. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29**, 353–366 (2001).
- Molyneaux, B. J., Arlotta, P., Hirata, T., Hibi, M. & Macklis, J. D. Fez1 is required for the birth and specification of corticospinal motor neurons. *Neuron* **47**, 817–831 (2005).
- Chen, B., Schaevelitz, L. & McConnell, S. Fez1 regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. *Proc. Natl Acad. Sci. USA* **102**, 17184–17189 (2005).
- Chen, J., Rasin, M., Kwan, K. & Sestan, N. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. *Proc. Natl Acad. Sci. USA* **102**, 17792–17797 (2005).
- Ariani, F. *et al.* FOXP1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **83**, 89–93 (2008).
- Kwan, K. *et al.* SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proc. Natl Acad. Sci. USA* **105**, 16021–16026 (2008).
- Knob, R. *et al.* Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years. *PLoS ONE* **5**, e8809 (2010).
- Han, W. *et al.* TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract. *Proc. Natl Acad. Sci. USA* **108**, 3041–3046 (2011).
- McKenna, W. L. *et al.* Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development. *J. Neurosci.* **31**, 549–564 (2011).
- Perroud, N. *et al.* Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J.* advance online publication, (<http://dx.doi.org/10.1038/tpj.2010.70>) (2010).
- Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460**, 744–747 (2009).
- Petanjek, Z., Judas, M., Kostović, I. & Uylings, H. Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern. *Cereb. Cortex* **18**, 915–929 (2008).
- Huttenlocher, P. R. & Dabholkar, A. S. Regional differences in synaptogenesis in human cerebral cortex. *J. Comp. Neurol.* **387**, 167–178 (1997).
- Stranger, B. E. *et al.* Population genomics of human gene expression. *Nature Genet.* **39**, 1217–1224 (2007).
- Heinzen, E. L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008).
- Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol. Psychiatry* **15**, 779–784 (2010).
- Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
- Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nature Genet.* **39**, 1494–1499 (2007).
- Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445–458 (2009).
- Ren, C., Ren, C. H., Li, L., Goltsov, A. A. & Thompson, T. C. Identification and characterization of RTVP1/GLIPR1-like genes, a novel p53 target gene cluster. *Genomics* **88**, 163–172 (2006).
- Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* doi:10.1038/nature10524 (this issue).

Supplementary Information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank A. Belanger, V. Imamovic, R. Johnson, P. Larton, S. Lindsay, B. Poulos, J. Rajan, D. Rimm and R. Zielke for assistance with tissue acquisition, D. Singh for technical assistance, I. Kostovic and Z. Petanjek for dendrite measurements, P. Levitt for suggesting the inclusion of ITC in the study, and D. Karolchik and A. Zweig for help in creating tracks for the UCSC Genome Browser. We also thank A. Beckel-Mitchener, M. Freund, M. Gerstein, D. Geschwind, T. Insel, M. Judas, J. Knowles, E. Lein, P. Levitt, N. Parikshak and members of the Sestan laboratory for discussions and criticism. Tissue was obtained from several sources including the Human Fetal Tissue Repository at the Albert Einstein College of Medicine, the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Laboratory of Developmental Biology at the University of Washington (supported by grant HD000836 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development) and the Joint MRC/Wellcome Trust Human Developmental Biology Resource (<http://hdb.org>) at the IGH, Newcastle Upon Tyne (UK funding awards G0700089 and G082557). Support for predoctoral fellowships was provided by the China Scholarship Council (Y.Z.), the Portuguese Foundation for Science and Technology (A.M.M.S.), the Samsung Scholarship Foundation (Y.S.), a Fellowship of the German Academic Exchange Service – DAAD (S. Mayer) and NIDA grant DA026119 (T.G.). This work was supported by grants from the US National Institutes of Health (MH081896, MH089929, NS054273), the Kavli Foundation and NARSAD, and by a James S. McDonnell Foundation Scholar Award (N.S.).

**Author Contributions** H.J.K., Y.I.K., A.M.M.S., M.P., K.A.M., G.S., Y.S., M.B.J., Z.K., S. Mayer, S.F., S.U., S. Mane and N.S. performed and analysed the experiments. F.C., Y.Z., X.X., M.L., T.G. and M.R. analysed the data. Z.K., A.M.M.S., M.P., G.S., S.N.L., A.V., D.R.W., T.M.H., A.H., J.E.K. and N.S. participated in tissue procurement and examination. N.S. designed the study and wrote the manuscript, which all authors commented and edited.

**Author Information** Exon array have been deposited in the NCBI Gene Expression Omnibus under accession number GSE25219 and genotyping data have been deposited in the NCBI database of Genotypes and Phenotypes under accession number phs000406.v1.p1. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to N.S. ([nenad.sestan@yale.edu](mailto:nenad.sestan@yale.edu)).

# Kelvin–Helmholtz instabilities as the source of inhomogeneous mixing in nova explosions

Jordi Casanova<sup>1,2</sup>, Jordi José<sup>1,2</sup>, Enrique García-Berro<sup>3,2</sup>, Steven N. Shore<sup>4</sup> & Alan C. Calder<sup>5</sup>

Classical novae<sup>1,2</sup> are thermonuclear explosions in binary stellar systems containing a white dwarf accreting material from a close companion star. They repeatedly eject  $10^{-4}$ – $10^{-5}$  solar masses of nucleosynthetically enriched gas into the interstellar medium, recurring on intervals of decades to tens of millennia. They are probably the main sources<sup>3,4</sup> of Galactic  $^{15}\text{N}$ ,  $^{17}\text{O}$  and  $^{13}\text{C}$ . The origin of the large enhancements and inhomogeneous distribution of these species observed in high-resolution spectra<sup>5</sup> of ejected nova shells has, however, remained unexplained for almost half a century<sup>6</sup>. Several mechanisms<sup>7</sup>, including mixing by diffusion<sup>8</sup>, shear<sup>9</sup> or resonant gravity waves<sup>10</sup>, have been proposed in the framework of one-dimensional or two-dimensional simulations, but none has hitherto proven successful because convective mixing can only be modelled accurately in three dimensions. Here we report the results of a three-dimensional nuclear-hydrodynamic simulation of mixing at the core–envelope interface during nova outbursts. We show that buoyant fingering drives vortices from the Kelvin–Helmholtz instability, which inevitably enriches the accreted envelope with material from the outer white-dwarf core. Such mixing also naturally produces large-scale chemical inhomogeneities. Both the metallicity enhancement and the intrinsic dispersions in the abundances are consistent with the observed values.

High-resolution spectra of nova shells, taken before the ejecta have undergone substantial modifications through interactions with the stellar companions<sup>11,12</sup> or with interstellar medium, always show highly fragmented, chemically enriched and inhomogeneous shells. V1974 Cyg, for instance, although a neon nova, showed large—more than threefold—differences in the C/He abundance ratios between two knots resolved in the spectrum a few years after outburst<sup>13</sup>. A similar abundance pattern has also been observed in many other novae, such as HR Del 1967 and DQ Her 1934, for which the shells are spatially resolved. Comparison of the infrared spectrum with the ultraviolet spectrum for the same ions at stages when the former, then the latter, turn optically thin shows the same structures to be present even during the previous opaque stages<sup>14</sup>, as do the multiple line systems known for decades from optical spectra<sup>1,15</sup>.

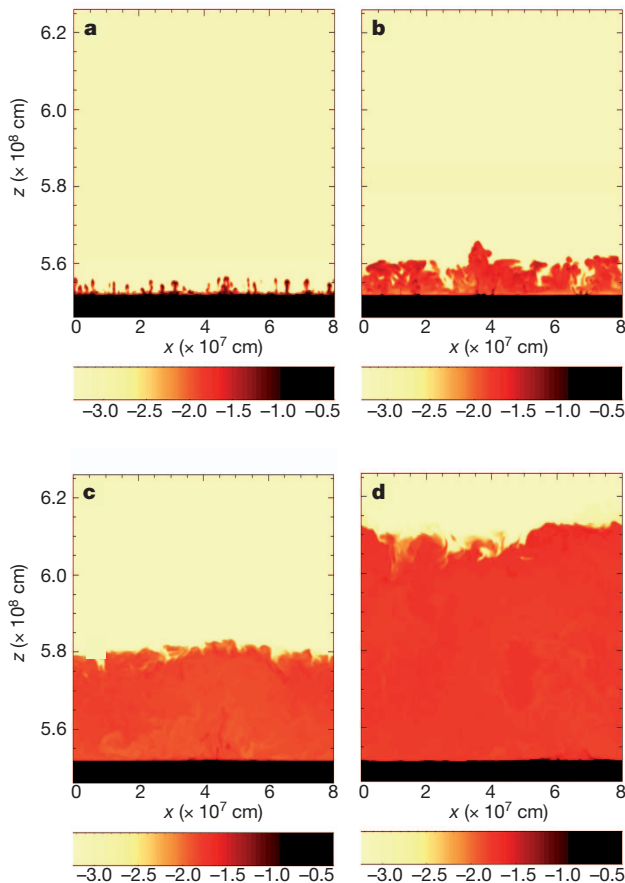
The nova outbursts (whether involving He-rich, CO-rich or ONe-rich white dwarfs) are triggered by nuclear processes dominated by CNO-cycle reactions that produce  $^{13}\text{N}$ ,  $^{14,15}\text{O}$  and  $^{17}\text{F}$  far in excess of solar abundances. Convection begins as soon as the temperature gradient becomes super-adiabatic, powered by the energy released from nuclear reactions (driven mainly by proton captures and  $\beta^+$  decays<sup>3</sup>, with a main nuclear path running close to the valley of stability), and is critical to the explosion, transferring a fraction of these abundant, short-lived species to the outer envelope layers. The energy released when these nuclei decay lifts degeneracy and drives the expansion, and ultimate ejection, of the polluted strata<sup>16</sup>. Hence, any attempt to explain the peculiarities in the abundance pattern observed in nova ejecta requires an accurate model of the physical processes that occur

during the explosion, namely the nuclear processes during the thermonuclear runaway along with the convective mass and energy transport. The peak temperatures reached during a nova explosion are constrained by the chemical abundance pattern inferred from the ejecta and do not exceed  $4 \times 10^8$  K, so it is unlikely that the observed metallicity enhancements can be due to thermonuclear processes driven by CNO breakout. Instead, mixing at the core–envelope interface is the more likely explanation. This cannot be modelled in the one-dimensional framework traditionally used in nova nucleosynthesis simulations, because mixing is inhibited by the one-dimensional mixing-length formalism of convection.

Early attempts in two-dimensional simulations<sup>17–23</sup> have shown that the onset of convection at the late stages of the thermonuclear runaway, driven by shear flows at the core–envelope interface, will ultimately dredge up chemically enriched material into the envelope. However, two-dimensional approximations for convection are unrealistic<sup>24,25</sup>; the conservation of vorticity imposed by the two-dimensional geometry forces the small convective cells to merge into large eddies, with a size comparable to the pressure scale height of the envelope. In contrast, in three-dimensional fully developed turbulent convection, eddies will become unstable and consequently will break up and filament, transferring their energy to progressively smaller scales<sup>26,27</sup>. These structures, vortices and filaments, must undergo a similar fate down to roughly the Kolmogorov scale,  $\eta = (\nu^3/\varepsilon)^{1/4}$ , where  $\nu$  is the kinematic viscosity and  $\varepsilon$  is the energy dissipation rate. Until now there has been no indication of such a cascade in nova hydrodynamic simulations. We show that with sufficient resolution, proper treatment of the nuclear processing, and long time spans, this turbulent energy transfer occurs and solves the mixing problem.

We therefore performed three-dimensional simulations of mixing at the core–envelope interface during classical nova explosions with the multidimensional, Eulerian, explicit code FLASH. The initial model<sup>17</sup> consists of a  $1M_{\odot}$  CO white dwarf that accretes solar composition matter ( $Z = 0.02$ ) at a rate of  $5 \times 10^{-9} M_{\odot} \text{ yr}^{-1}$ . The model was evolved in one dimension and subsequently mapped onto a three-dimensional Cartesian grid of  $800 \times 800 \times 800 \text{ km}^3$ , when the temperature at the base of the envelope reached  $10^8$  K (see Fig. 1 and Supplementary Information). The model was relaxed to guarantee hydrostatic equilibrium. A top-hat 2-km-wide temperature perturbation<sup>23</sup> (5% amplitude) was imposed close to the core–envelope interface. As tested thoroughly in two dimensions, the specific choice of the initial perturbation (that is, duration, strength, location and size), the resolution adopted or the size of the computational domain does not have a severe impact on the results of the simulation. The initial perturbation drives a shear flow triggering Kelvin–Helmholtz instabilities about 150 s after the start of the simulation. Small convection cells develop as soon as material is dredged up into the envelope. Moreover, the fluid velocity remains below the speed of sound, confirming that a nova outburst is driven by a (subsonic) deflagration rather than a

<sup>1</sup>Departament de Física i Enginyeria Nuclear, EUETIB, Universitat Politècnica de Catalunya, C./ Comte d'Urgell 187, E-08036 Barcelona, Spain. <sup>2</sup>Institut d'Estudis Espacials de Catalunya, Ed. Nexus-201, C./ Gran Capità 2-4, E-08034 Barcelona, Spain. <sup>3</sup>Departament de Física Aplicada, Universitat Politècnica de Catalunya, C./ Esteve Terrades 5, E-08860 Castelldefels (Barcelona), Spain. <sup>4</sup>Dipartimento di Fisica 'Enrico Fermi', Università di Pisa and Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Largo B. Pontecorvo 3, I-56127 Pisa, Italy. <sup>5</sup>Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA.



**Figure 1 | Mixing driven by Kelvin–Helmholtz instabilities.** Snapshots of the development of Kelvin–Helmholtz instabilities at  $t = 151$  s (a), 193 s (b), 296 s (c) and 379 s (d), shown in terms of the  $^{12}\text{C}$  mass fraction on a logarithmic scale. Dredging of core material driven by Kelvin–Helmholtz instabilities translates into a mass-averaged abundance of CNO nuclei in the envelope of 0.118, 0.129, 0.157 and 0.182, respectively. The mean CNO abundance at the end of the simulation reaches 0.20 by mass (see Supplementary Movies). Calculations were performed with FLASH<sup>29</sup>, a parallelized explicit Eulerian code, based on the piecewise parabolic interpolation of physical quantities for solving the hydrodynamic equations, and with adaptive mesh refinement (with four or five levels of refinement). Simulations were run at the MareNostrum supercomputer, requiring 150,000 CPU hours with 256 (occasionally 512) processors. The typical resolution adopted was  $3.12 \times 3.12 \times 3.12 \text{ km}^3$ , with a maximum resolution of  $1.56 \times 1.56 \times 1.56 \text{ km}^3$ . The three-dimensional computational domain initially comprised 112 radial layers—including the outermost part of the CO core—and 512 horizontal layers along both horizontal axes. The mass of the accreted envelope was about  $2 \times 10^{-5} M_{\odot}$ . Nuclear energy generation was handled through a network of 13 species ( $^1\text{H}$ ,  $^4\text{He}$ ,  $^{12,13}\text{C}$ ,  $^{13,14,15}\text{N}$ ,  $^{14,15,16,17}\text{O}$  and  $^{17,18}\text{F}$ ), connected through 18 nuclear processes. Periodic conditions were imposed at the four vertical boundaries of the computational domain, and hydrostatic equilibrium with an outflow constraint at the top and a reflecting constraint at the bottom was imposed on the velocity at the horizontal boundaries<sup>30</sup>. Other details on the input physics are identical to those adopted in earlier two-dimensional simulations<sup>22,23</sup>.

(supersonic) detonation. Even for a point-like ignition, the burning front quickly spreads horizontally, such that the expansion and progress of the explosion proceed in almost spherical symmetry. This confirms early estimates<sup>28</sup> of the velocity of the deflagration front spreading through the stellar surface, in the form  $v_{\text{def}} \sim (h_p v_{\text{conv}} / \tau_{\text{burn}})^{1/2}$ , where  $h_p$  is the pressure scale height,  $v_{\text{conv}}$  the characteristic convective velocity and  $\tau_{\text{burn}}$  the characteristic timescale for fuel burning. Typical values for nova outbursts yield  $v_{\text{def}} = 10^4 \text{ cm s}^{-1}$ ; that is, a flame propagating halfway across the stellar surface in about 1.3 days.

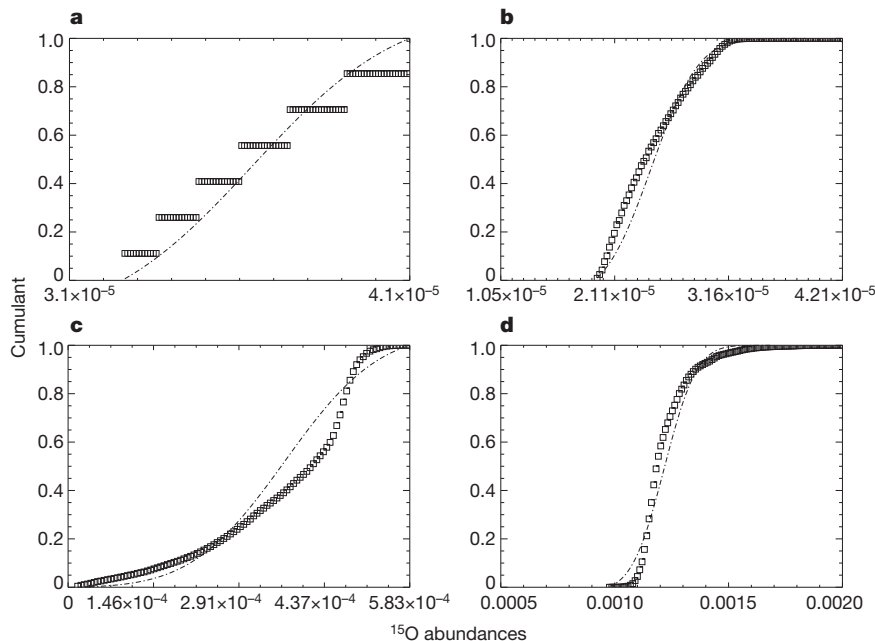
At  $t = 400$  s, matter crosses the outer computational boundary and we stop the calculations because of the Eulerian nature of the FLASH

code. At this stage, the envelope base has reached a peak temperature of  $1.82 \times 10^8 \text{ K}$ , and the mean (mass-averaged) metallicity in the envelope reaches  $Z \approx 0.20$ . This agrees with observations of CO novae and with previous two-dimensional modelling<sup>22,23</sup>, suggesting that the dimensionality of the convective treatment is not crucial to the mean values but is important for the details (in particular for an accurate description of convective transport). The simulations reveal that these last stages are, however, intermittent, and our three-dimensional simulations seem to resolve at least the upper dissipation range (see Supplementary Movies). In that sense, the Kolmogorov scale,  $\eta$ , provides an estimate of the size of the smallest eddies present in the flow. Also at this stage, the Reynolds number becomes sufficiently small and molecular viscosity is effective in dissipating the kinetic energy into heat.

In our simulations, the burning advances along with the development of persistent density contrasts of large size, comparable to the thickness of the layer and much larger than the burning transition zone. These become turbulent, and the models require a fully three-dimensional treatment to capture the full spectrum of the plumes and vortex structures. The resulting abundances from our simulations have another particular feature that agrees with and explains the observations: the structures are chemically inhomogeneous. This is a relic of Rayleigh–Taylor instabilities that grow during the initial stages of the ejection. These structures appear within the burning zone (that is, the initial plumes) and drive a secondary Kelvin–Helmholtz vortex cascade that induces turbulent motions. In this regime of very high Reynolds number, the fluid motion is extremely complex and develops structures on all scales. In particular, vortices and filaments appear that are a signature of the intermittency in classical shear-flow turbulence. These are more evident in two-dimensional simulations, but they cascade rapidly into smaller eddies and filaments, subject to recombination and extension, as the burning continues. Such dissipation is intermittent, as predicted in the Kolmogorov–Obukov theory of turbulence<sup>26,27</sup>, generating coherent, persisting structures that advect with the expanding layers. Because the nuclear reaction rates are density sensitive, higher-density knots have a different nuclear history from that of the background, and this is best characterized by the abundance distribution function behind the deflagration.

This can be clearly seen in Fig. 2, which shows a sample of the time evolution of the cumulant function. The initial abundance for  $^{15}\text{O}$ , our trace species (whose abundance is increased by the deep non-uniform mixing after the onset of thermal buoyant turbulence), is a narrow initial distribution that evolves into a stable form with a lower cutoff and a power-law tail towards high abundances. Unlike the single value obtained in one-dimensional models, we find a 30% dispersion in the main component, fitted by a Gaussian (the dashed line in Fig. 2), and containing about 10% of the total volume, an extended, non-Gaussian ‘fat tail’ whose maximum abundance (at the 1% level) extends up to  $13\sigma$  from the mean  $^{15}\text{O}$  for the volume. Multiwavelength spectroscopic analyses during the nova nebular stage, when the ejecta are essentially transparent, frequently find large dispersions from line to line in the abundance ratios of the principal chemical species. Although usually assumed to result from measurement uncertainties, which are frequently lower than the derived dispersions, this may instead be a physically significant result: a signature of the turbulence generated during the thermonuclear runaway.

Spatially resolved knots in several classical novae, notably V1974 Cyg (Nova Cyg 1992) demonstrate the reality of this inhomogeneity in the ejecta, up to threefold, long before the mixtures are diluted by interaction with the interstellar medium. These were first seen in the first weeks of the expansion of the ejecta across different spectral regions, from the infrared to the ultraviolet, and were invariant in velocity and contrast as each layer was exposed, after the opacity drop in the expanding medium. The contrast between these residual structures will be further amplified by the supersonic motions that follow the stage shown here. Although some structuring may result from collision of the ejecta with the accretion disk around the white dwarf and with the



**Figure 2 | Cumulant distributions.** Cumulant distribution function for  $^{15}\text{O}$  (squares), at  $t = 190$  s (a), 240 s (b), 315 s (c) and 397 s (d), compared with a Gaussian process (dashed line) with the same mean and dispersion. The cumulant distribution shows the probability that a variable  $x$  has a value equal to or lower than a threshold value  $a$  (that is,  $P[x \leq a]$ ). The cumulant distribution is estimated by means of an arbitrary layer 100 km thick ( $5.75 \times 10^8$  to  $5.85 \times 10^8$  cm, with 27,648 data points), a size comparable with the pressure scale height. The evolution shows the effects of the intermittent bursts seen in the Supplementary Movies. The ranges predicted are in agreement with those inferred from early-stage spectroscopy of the ejecta. These can be compared directly with velocity-dependent line profile ratios for different ions and elements. The distributed long-lived  $^{15}\text{O}$ -enriched eddies and filaments, when they undergo  $\beta$ -decay, provide distributed but distinct trigger sites for the final explosive ejection of the outer layers on a spatial scale consistent with the inferred sizes of the knots based on high-resolution line profiles.

companion, the universality of the filamentation—independently of the parameters of the binary system and the characteristic times at which they are observed—points back to the first moments of the explosion. Many spatially resolved nova shells, notably GK Per 1901, HR Del 1967 and the recurrent nova T Pyx, provide ample evidence of the complexity of the filaments in late stages of the expansion.

These fluctuations in abundance, density and, in turn, opacity will leave imprints not only in the ejecta but also in several aspects accompanying nova outbursts. This will certainly influence the modelling of the initial nova outflows, of dust formation (which will condense more rapidly in the more  $^{12}\text{C}$ -enriched regions that survive within the ejecta) and of the maximum achievable luminosities during the explosion (for instance the extent to which classical nova explosions can exceed the Eddington luminosity and where this occurs). In addition, the long-standing problem of the high masses accumulated before the explosion and the subsequent discrepancy between the observed and computed ejecta masses is naturally resolved by deep mixing of the envelope material<sup>4,6</sup>. A persistent problem in one-dimensional models has been the need for some dynamical premixing, either by shear instability or artificially enhanced convection, to match the observed abundance pattern; three-dimensional models, as shown in this paper, require neither.

Received 26 April; accepted 30 August 2011.

Published online 19 October 2011.

1. Bode, M. F. & Evans, A. (eds). *Classical Novae* 2nd edn (Cambridge Univ. Press, 2008).
2. Hernanz, M. & José, J. (eds). *Classical Nova Explosions* (AIP, 2002).
3. José, J., Hernanz, M. & Iliadis, C. Nucleosynthesis in classical novae. *Nucl. Phys. A* **777**, 550–578 (2006).
4. Starrfield, S., Iliadis, C. & Hix, W. R. in *Classical Novae* 2nd edn (eds Bode, M. F. & Evans, A.) 77–101 (Cambridge Univ. Press, 2008).
5. Gehrz, R. D., Truran, J. W., Williams, R. E. & Starrfield, S. Nucleosynthesis in classical novae and its contribution to the interstellar medium. *Publ. Astron. Soc. Pacif.* **110**, 3–26 (1998).
6. José, J. & Shore, S. N. in *Classical Novae* 2nd edn (eds Bode, M. F. & Evans, A.) 121–140 (Cambridge Univ. Press, 2008).
7. Shore, S. N., Livio, M. & van den Heuvel, E. P. J. *Interacting Binaries* (Springer, 1994).
8. Prialnik, D. & Kovetz, A. The effect of diffusion on prenova evolution—CNO-enriched envelopes. *Astrophys. J.* **281**, 367–374 (1984).
9. Kutter, G. S. & Sparks, W. M. Stellar accretion of matter possessing angular momentum. *Astrophys. J.* **321**, 386–393 (1987).
10. Rosner, R., Alexakis, A., Young, Y.-N., Truran, J. W. & Hillebrandt, W. On the C/O enrichment of nova ejecta. *Astrophys. J.* **562**, L177–L179 (2001).
11. Porter, J. M., O'Brien, T. J. & Bode, M. F. On the asphericity of nova remnants caused by rotating white dwarf envelopes. *Mon. Not. R. Astron. Soc.* **296**, 943–948 (1998).
12. Vaytet, N. M. H., O'Brien, T. J. & Rushton, A. P. Evidence for ablated flows in the shell of the nova DQ Herculis. *Mon. Not. R. Astron. Soc.* **380**, 175–180 (2007).

13. Shore, S. N., Starrfield, S., Ake, T. B., III & Hauschildt, P. H. Spatially resolved spectra of V1974 Cygni (Nova Cygni 1992) with the Goddard High Resolution Spectrograph. *Astrophys. J.* **490**, 393–400 (1997).
14. Gehrz, R. D. in *Classical Novae* 2nd edn (eds Bode, M. F. & Evans, A.) 167–193 (Cambridge Univ. Press, 2008).
15. McLaughlin, D. B. The behaviour of absorption systems in spectra of novae. *Ann. Astrophys.* **27**, 450–461 (1964).
16. Starrfield, S., Truran, J. W., Sparks, W. M. & Kutter, G. S. CNO abundances and hydrodynamic models of the nova outburst. *Astrophys. J.* **176**, 169–176 (1972).
17. Glasner, S. A. & Livne, E. Convective hydrogen burning down a nova outburst. *Astrophys. J.* **445**, L149–L151 (1995).
18. Glasner, S. A., Livne, E. & Truran, J. W. Reactive flow in nova outbursts. *Astrophys. J.* **475**, 754–762 (1997).
19. Kerckel, A., Hillebrandt, W. & Truran, J. W. Two-dimensional simulations of the thermonuclear runaway in an accreted atmosphere of a C+O white dwarf. *Astron. Astrophys.* **337**, 379–392 (1998).
20. Glasner, S. A., Livne, E. & Truran, J. W. The sensitivity of multidimensional nova calculations to the outer boundary condition. *Astrophys. J.* **625**, 347–350 (2005).
21. Glasner, S. A., Livne, E. & Truran, J. W. Novae: the evolution from onset of convection to the runaway. *Astrophys. J.* **665**, 1321–1333 (2007).
22. Casanova, J., José, J., García-Berro, E., Calder, A. & Shore, S. N. On mixing at the core–envelope interface during classical nova outbursts. *Astron. Astrophys.* **513**, L5 (2010).
23. Casanova, J., José, J., García-Berro, E., Calder, A. & Shore, S. N. Mixing in classical novae: a 2-D sensitivity study. *Astron. Astrophys.* **527**, A5 (2011).
24. Kerckel, A., Hillebrandt, W. & Truran, J. W. Three-dimensional simulations of classical novae. *Astron. Astrophys.* **345**, 831–840 (1999).
25. Arnett, D., Meakin, C. & Young, P. A. Turbulent convection in stellar interiors. II. The velocity field. *Astrophys. J.* **690**, 1715–1729 (2009).
26. Pope, S. B. *Turbulent Flows* (Cambridge Univ. Press, 2000).
27. Shore, S. N. *Astrophysical Hydrodynamics: An Introduction* (Wiley, 2007).
28. Fryxell, B. A. & Woosley, S. E. Finite propagation time in multidimensional thermonuclear runaways. *Astrophys. J.* **261**, 332–336 (1982).
29. Fryxell, B. et al. FLASH: an adaptive mesh hydrodynamics code for modeling astrophysical thermonuclear flashes. *Astrophys. J.* **131** (Suppl.), 273–334 (2000).
30. Zingale, M. et al. Mapping initial hydrostatic models in Godunov codes. *Astrophys. J.* **143** (Suppl.), 539–565 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The software used in this work was developed in part by the Department of Energy-supported Alliances Center for Astrophysical Thermonuclear Flashes at the University of Chicago. This work was partly supported by Spanish Ministerio de Educación y Ciencia grants, by the Agència de Gestió d'Ajuts Universitaris i de Recerca de Generalitat de Catalunya, by the E.U. European Fund for Regional Development, and by the European Science Foundation EUROCORES Program EuroGENESIS. We also acknowledge the Barcelona Supercomputing Center for allocation of time at the MareNostrum supercomputer.

**Author Contributions** All authors contributed equally to the results presented here.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.J. (jordi.jose@upc.edu).

# A Pluto-like radius and a high albedo for the dwarf planet Eris from an occultation

B. Sicardy<sup>1,2,3</sup>, J. L. Ortiz<sup>4</sup>, M. Assafin<sup>5</sup>, E. Jehin<sup>6</sup>, A. Maury<sup>7</sup>, E. Lellouch<sup>1</sup>, R. Gil Hutton<sup>8</sup>, F. Braga-Ribas<sup>1,9</sup>, F. Colas<sup>10</sup>, D. Hestroffer<sup>10</sup>, J. Lecacheux<sup>1</sup>, F. Roques<sup>1</sup>, P. Santos-Sanz<sup>1</sup>, T. Widemann<sup>1</sup>, N. Morales<sup>4</sup>, R. Duffard<sup>4</sup>, A. Thirouin<sup>4</sup>, A. J. Castro-Tirado<sup>4</sup>, M. Jelínek<sup>4</sup>, P. Kubánek<sup>4</sup>, A. Sota<sup>4</sup>, R. Sánchez-Ramírez<sup>4</sup>, A. H. Andrei<sup>5,9</sup>, J. I. B. Camargo<sup>5,9</sup>, D. N. da Silva Neto<sup>9,11</sup>, A. Ramos Gomes Jr<sup>5</sup>, R. Vieira Martins<sup>5,9,10</sup>, M. Gillon<sup>6</sup>, J. Manfroid<sup>6</sup>, G. P. Tozzi<sup>12</sup>, C. Harlinton<sup>13</sup>, S. Saravia<sup>7</sup>, R. Behrend<sup>14</sup>, S. Mottola<sup>15</sup>, E. García Melendo<sup>16,17</sup>, V. Peris<sup>18</sup>, J. Fabregat<sup>18</sup>, J. M. Madiedo<sup>19</sup>, L. Cuesta<sup>20</sup>, M. T. Eibe<sup>20</sup>, A. Ullán<sup>20</sup>, F. Organero<sup>21</sup>, S. Pastor<sup>22</sup>, J. A. de los Reyes<sup>22</sup>, S. Pedraz<sup>23</sup>, A. Castro<sup>24</sup>, I. de la Cueva<sup>25</sup>, G. Muler<sup>26</sup>, I. A. Steele<sup>27</sup>, M. Cebrián<sup>28</sup>, P. Montañés-Rodríguez<sup>28</sup>, A. Oscoz<sup>28</sup>, D. Weaver<sup>29</sup>, C. Jacques<sup>30</sup>, W. J. B. Corradi<sup>31</sup>, F. P. Santos<sup>31</sup>, W. Reis<sup>31</sup>, A. Milone<sup>32</sup>, M. Emilio<sup>33</sup>, L. Gutiérrez<sup>34</sup>, R. Vázquez<sup>34</sup> & H. Hernández-Toledo<sup>35</sup>

**The dwarf planet Eris is a trans-Neptunian object with an orbital eccentricity of 0.44, an inclination of 44 degrees and a surface composition very similar to that of Pluto<sup>1</sup>. It resides at present at 95.7 astronomical units (1 AU is the Earth-Sun distance) from Earth, near its aphelion and more than three times farther than Pluto. Owing to this great distance, measuring its size or detecting a putative atmosphere is difficult. Here we report the observation of a multi-chord stellar occultation by Eris on 6 November 2010 UT. The event is consistent with a spherical shape for Eris, with radius  $1,163 \pm 6$  kilometres, density  $2.52 \pm 0.05$  grams per cm<sup>3</sup> and a high visible geometric albedo,  $p_v = 0.96^{+0.09}_{-0.04}$ . No nitrogen, argon or methane atmospheres are detected with surface pressure larger than  $\sim 1$  nanobar, about 10,000 times more tenuous than Pluto's present atmosphere<sup>2–5</sup>. As Pluto's radius is estimated<sup>3–8</sup> to be between 1,150 and 1,200 kilometres, Eris appears as a Pluto twin, with a bright surface possibly caused by a collapsed atmosphere, owing to its cold environment. We anticipate that this atmosphere may periodically sublimate as Eris approaches its perihelion, at 37.8 astronomical units from the Sun.**

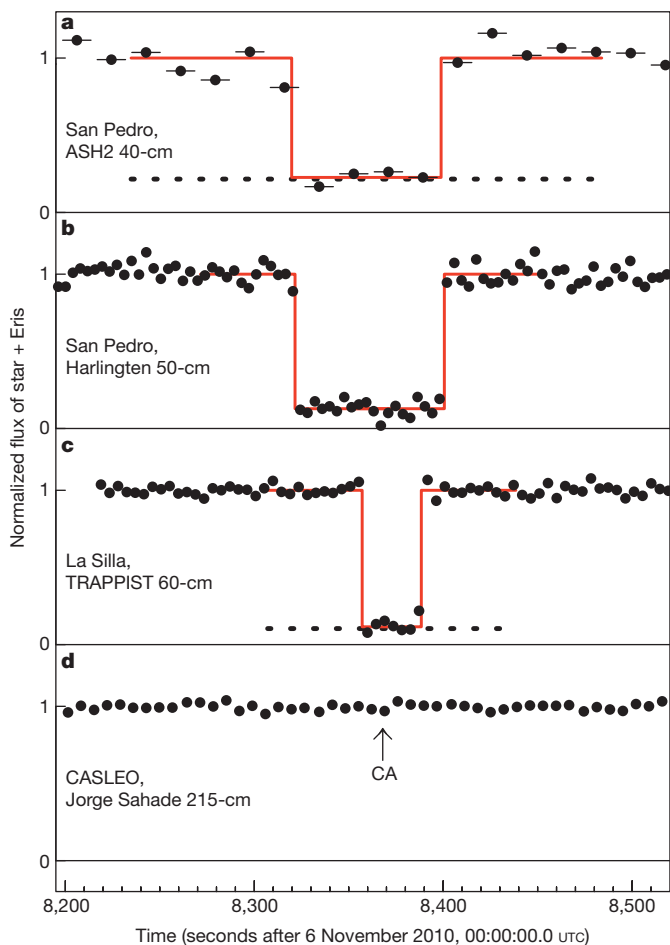
The dwarf planet (136199) Eris was discovered<sup>9</sup> in 2005. Its radius has been estimated to be  $1,200 \pm 100$  km on the basis of direct imaging<sup>10</sup>, although detection of its thermal flux provided another estimate<sup>11</sup> of  $1,500 \pm 200$  km, potentially making it larger even than Pluto, and the largest known dwarf planet. The motion of Dysnomia (Eris' satellite) provides Eris' mass,  $M_E = (1.66 \pm 0.02) \times 10^{22}$  kg, 27% larger than Pluto's mass<sup>12</sup>. No short-term (day-scale) brightness variability has been detected for Eris at the 1% level<sup>13,14</sup>, suggesting either a spherical body with no albedo variegation, or—if elongated—a finely-tuned, pole-on viewing geometry. The spectrum of Eris is very similar to that of Pluto and reveals a methane-ice-rich cover, and another dominant ice, presumably nitrogen, but not excluding argon<sup>1</sup>.

Stellar occultations by Eris are rare, as it subtends a minuscule angular diameter ( $\sim 0.03$  arcsec) while currently moving in severely depleted stellar fields at an angular rate of  $\sim 1.5$  arcsec h<sup>-1</sup> at most. Using the techniques described in ref. 15, we predicted one Eris occultation in 2010, on November 6 UT. We attempted observations from 26 stations, and the occultation was detected from two sites in Chile, with two detections at San Pedro de Atacama (San Pedro for short) with the Harlinton and ASH2 telescopes, 20 m from each other, and one detection at La Silla, with the TRAPPIST telescope (for details, see Fig. 1, Supplementary Figs 1 and 2, and Supplementary Tables 1 and 2). Another station further south at Complejo Astronómico El Leoncito (CASLEO), Argentina, provided a light curve without occultation, but went close to Eris' shadow edge ( $\sim 200$  km; see Fig. 2).

The San Pedro and La Silla observations provide two occultation segments—or 'chords'—whose four extremities are used to constrain Eris' size (red segments in Fig. 2). When deriving the occultation times, it appeared that two equally satisfactory solutions for the star reappearance time at the Harlinton telescope in San Pedro are possible, yielding two different chord lengths. These two solutions are separated by 1.2 s, and are respectively called solution 1 and solution 2, in chronological order. This ambiguity is due to the fact that the star reappearance occurred during a gap between consecutive exposures, corresponding to a net loss of information. The ASH2 data collected next to Harlinton did not provide enough signal-to-noise ratio to discriminate between these two solutions, and are not used in the fit described below (see Supplementary Information). As a dwarf planet, Eris is expected to be in hydrostatic equilibrium under gravity and centrifugal forces. The most general apparent limb shape is then an ellipse with semi-axes  $a' > b'$  with effective radius  $R_E = \sqrt{a'b'}$ , defined as the radius of a disk that has the same apparent surface area as the actual body. This shape stems either from an oblate Maclaurin spheroid

<sup>1</sup>LESIA-Observatoire de Paris, CNRS, Université Pierre et Marie Curie, Université Paris-Diderot, 11, Rue Marcelin Berthelot, 92195 Meudon cedex, France. <sup>2</sup>Université Pierre et Marie Curie, 4, Place Jussieu, 75252 Paris cedex 5, France. <sup>3</sup>Institut Universitaire de France, 103, Boulevard Saint Michel, 75005 Paris, France. <sup>4</sup>Instituto de Astrofísica de Andalucía, CSIC, Apartado 3004, 18080 Granada, Spain.

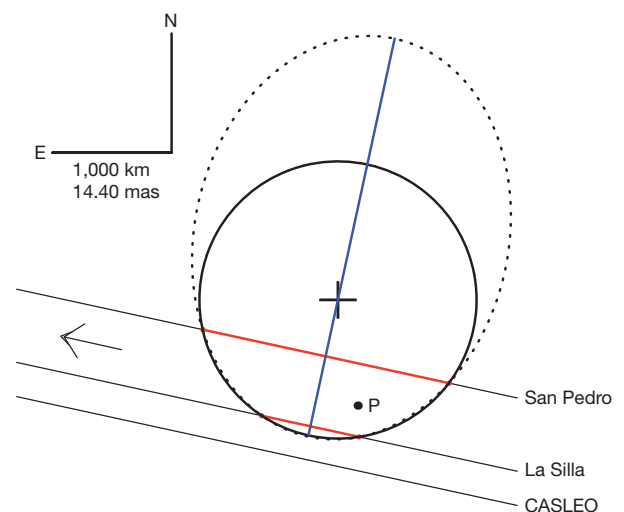
<sup>5</sup>Observatório do Valongo/UFRJ, Ladeira Pedro Antonio 43, CEP 20.080-090 Rio de Janeiro, RJ, Brazil. <sup>6</sup>Institut d'Astrophysique de l'Université de Liège, Allée du 6 Août 17, B-4000 Liège, Belgium. <sup>7</sup>San Pedro de Atacama Celestial Explorations, Casilla 21, San Pedro de Atacama, Chile. <sup>8</sup>Complejo Astronómico El Leoncito (CASLEO) and San Juan National University, Avenida España 1512 sur, J5402DSP, San Juan, Argentina. <sup>9</sup>Observatório Nacional/MCT, Rua General José Cristino 77, CEP 20921-400 Rio de Janeiro, RJ, Brazil. <sup>10</sup>IMCCE-Observatoire de Paris, CNRS, Université Pierre et Marie Curie, Université Lille 1, 77 Avenue Denfert-Rochereau, 75014 Paris, France. <sup>11</sup>Centro Universitário Estadual da Zona Oeste, Avenida Manual Caldeira de Alvarenga 1203, CEP, 23.070-200 Rio de Janeiro, RJ, Brazil. <sup>12</sup>INAF, Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy. <sup>13</sup>Caisey Harlinton Observatory, The Grange, Scarrow Beck Road, Erpingham, Norfolk NR11 7QX, UK. <sup>14</sup>Observatoire de Genève, CH-1290 Sauverny, Switzerland. <sup>15</sup>DLR – German Aerospace Center, Rutherfordstrasse 2, 12489 Berlin, Germany. <sup>16</sup>Fundació Privada Observatori Esteve Duran, 08553 Seva, Spain. <sup>17</sup>Institut de Ciències de l'Espai (CSIC-IEEC), Facultat de Ciències, Torre C5, 08193 Bellaterra, Spain. <sup>18</sup>Observatori Astronòmic, Universitat de València, Calle Catedrático José Beltrán 2, 46980 Paterna, Spain. <sup>19</sup>Universidad de Huelva, Facultad de Ciencias Experimentales, Avenida de las Fuerzas Armadas S/N, 21071 Huelva, Spain. <sup>20</sup>Centro de Astrobiología (CSIC-INTA), Carretera de Ajalvir, km 4, 28850 Torrejón de Ardoz, Madrid, Spain. <sup>21</sup>Observatorio astronómico de La Hita, 45840 La Puebla de Almoradil (Toledo), Spain. <sup>22</sup>Observatorio de la Murta, 30153 Murcia, Spain. <sup>23</sup>Calar Alto Observatory, Centro Astronómico Hispano Alemán, Calle Jesús Durbán Remón, 2, 04004 Almería, Spain. <sup>24</sup>Sociedad Malagueña de Astronomía, Centro Cultural José María Gutiérrez Romero, Calle República Argentina 9, 29016 Málaga, Spain. <sup>25</sup>Astroimagen, Abad y Lasiera 58Bis, 07800 Ibiza, Spain. <sup>26</sup>Observatorio Nazaret, 35539 Nazaret, Lanzarote, Spain. <sup>27</sup>Liverpool JMU, Twelve Quays House, Egerton Wharf, Birkenhead CH41 1LD, UK. <sup>28</sup>Instituto de Astrofísica de Canarias, Vía Láctea s/n, 38205 La Laguna, Tenerife, Spain. <sup>29</sup>Observatório Astronômico Christus, Colégio Christus, Rua João Carvalho 630, Aldeota, CEP 60140140, Fortaleza, CE, Brazil. <sup>30</sup>Observatório CEAMIG-REA, CEP31545-120, Belo Horizonte, MG, Brazil. <sup>31</sup>Departamento de Física – Instituto de Ciências Exatas – Universidade Federal de Minas Gerais, Avenida Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil. <sup>32</sup>Instituto Nacional de Pesquisas Espaciais (INPE-MCT), Divisão de Astrofísica, Avenida dos Astronautas, 1758, São José dos Campos, 12227-010, SP, Brazil. <sup>33</sup>Universidade Estadual de Ponta Grossa, O.A. – DEGE, Avenida Carlos Cavalcanti 4748, Ponta Grossa 84030-900, PR, Brazil. <sup>34</sup>Instituto de Astronomía, Universidad Nacional Autónoma de México, Carretera Tijuana-Ensenada, km 103, 22860 Ensenada, BC, Mexico. <sup>35</sup>Instituto de Astronomía, Universidad Nacional Autónoma de México, Apartado Postal 70-264, 04510 México DF, Mexico.



**Figure 1 | Eris occultation light curves.** The plots (black filled circles) show the flux of the star plus Eris, normalized to unity outside the occultation, versus time. No filter was used at any of the telescopes. **a**, The light curve from the ASH2 40-cm telescope at San Pedro, using a SBIG STL-11000M CCD camera, with  $2 \times 2$  pixel binning and a sub-frame of  $11.24 \times 9.71$  arcmin ( $272 \times 235$  pixels). The horizontal error bars indicate the total time intervals associated with each point (15 s, while the cycle time was 18.32 s). Those bars are too small to be visible on the other data sets. **b**, The light curve from the Harlingen 50-cm telescope at San Pedro using an Apogee U42 CCD camera ( $2 \times 2$  pixel binning; sub-frame,  $2.67 \times 2.67$  arcmin, or  $100 \times 100$  pixels; integration/cycle times, 3 and 3.88 s). **c**, The light curve from the 60-cm TRAPPIST telescope at La Silla, using a FLI ProLine PL3041-BB CCD camera ( $2 \times 2$  pixel binning; sub-frame,  $3.25 \times 3.25$  arcmin, or  $150 \times 150$  pixels; integration/cycle times, 3 and 4.55 s). **d**, The light curve from the 215-cm Jorge Sahade telescope at CASLEO, using a Roper Scientific Versarray 1300B CCD camera ( $3 \times 3$  pixel binning; sub-frame,  $2.62 \times 3.50$  arcmin, or  $77 \times 103$  pixels; integration/cycle times, 4 and 7 s). The horizontal dashed lines at the bottom of **a** and **c** represent Eris' contribution to the flux, showing that the star completely disappeared during the event (Supplementary Information, section 2). The red lines are the best square-well models fitted to the events. We show in **b** solution 2 for the light curve (solution 1 being very close at that scale, Supplementary Fig. 3). The vertical arrow in **d** shows the time of closest approach (CA) to the shadow edge at CASLEO, at 8,368 s UT.

(small angular momentum regime) with assumed equator-on viewing, or an elongated triaxial Jacobi ellipsoid (large angular momentum regime) observed pole-on, as implied by the absence of brightness variations.

We have five free parameters to adjust:  $a'$ , the apparent flattening ( $(a' - b')/a'$ ), the ellipse position angle  $P$  in the sky plane, and the two coordinates of its centre,  $f_{\odot}, g_{\odot}$  (Supplementary Table 3). With four chord extremities, our observations allow for an infinity of limb solutions. However, as the two chords have almost the same median lines (Fig. 2), this strongly suggests that Eris' shape is indeed close to spherical,



**Figure 2 | Measuring Eris' size.** The three oblique solid lines show the star trajectories relative to Eris, as seen from San Pedro, La Silla and CASLEO, with the arrow pointing towards the direction of motion. The San Pedro and La Silla timings provide the lengths of the two occultation segments, or 'chords' (in red); see solution 2 in Supplementary Table 3. The median lines of the two red segments are separated by only 5 km and coincide at that scale with the blue line. Celestial north is up and east is left. Scale bars: 1,000 km and 14.40 mas (1 mas corresponds to 69.436 km at Eris). The solid circle has a radius  $R_E = 1,163$  km, and is our preferred solution for Eris' size and shape, with the cross marking the position of the centre. The dot near 'P' indicates Dysnomia's orbit pole direction<sup>12</sup> projected onto Eris' surface. The dotted curve is an elliptic limb fitted to our occultation chords, with semi-major and -minor axes of  $a' = 1,708$  and  $b' = 1,317$  km, respectively, that is, an apparent effective radius of  $R_E = 1,500$  km, the value of Eris' radius previously derived from thermal measurements<sup>11</sup>. The long axis of the ellipse should be perpendicular to the occultation chords to within  $\pm 2^\circ$  in order to match our data points. This has a low probability (2%) of occurring for a random limb orientation between  $0$  and  $180^\circ$ . Furthermore, the ellipse has an aspect ratio  $b/a = 0.771$  that would require a fast rotator (with a period of 4.4 h) observed pole-on to within  $18^\circ$  to suppress the rotational light curve<sup>13,14</sup>. This has also a low probability (5%) of occurring for a randomly distributed pole orientation, making the dotted limb solution unlikely.

unless very special geometries occurred (see below). Using a circular model with three free parameters ( $R_E = a', f_{\odot}, g_{\odot}$ ) and adopting solution 2, we obtain  $R_E = 1,163 \pm 6$  km ( $1\sigma$  formal error). The minimum  $\chi^2_{\text{pdf}} = 1.38$ , indicates a satisfactory fit to the data (Supplementary Table 3). Moreover, the r.m.s. radial residual of 2.1 km is fully consistent with our formal timing errors. We may not exclude, however, the possibility that random topographic features with amplitude approximately  $\pm 3$  km exist along the limb, which would result in a slightly larger error bar for Eris' radius,  $R_E = 1,163 \pm 9$  km (see Supplementary Information). Solution 1 provides  $R_E = 1,140$  km, but with a high value  $\chi^2_{\text{pdf}} = 30.7$  ( $5.5\sigma$  level), and radial residuals of  $+11$  km and  $-16$  km at the beginning and end of the San Pedro chord, respectively. Topographic features of this size are unlikely on such a large icy body. This indicates that the spherical assumption is not correct for solution 1, and explains why we do not provide a formal error bar for that value.

Allowing for a non-zero flattening of Eris' limb, we find an infinity of possible solutions by fixing the position angle  $P$  and semi-major axis  $a'$  at various values. If Eris' rotation axis and Dysnomia's orbital pole are aligned, we find values of  $R_E$  in the range 1,105–1,155 km, smaller than the value 1,163 km derived above. Relaxing the constraint on Eris' orientation, we find that elliptical limb models can satisfactorily fit the occultation chords in 68.3% of the cases ( $1\sigma$  level) for  $R_E$  in the range  $1,165 \pm 90$  km (Supplementary Fig. 4). However, as  $R_E$  departs from 1,165 km, the flattening must rapidly increase, requiring fast rotations which are not supported by observations<sup>13,14</sup>. The extreme case of

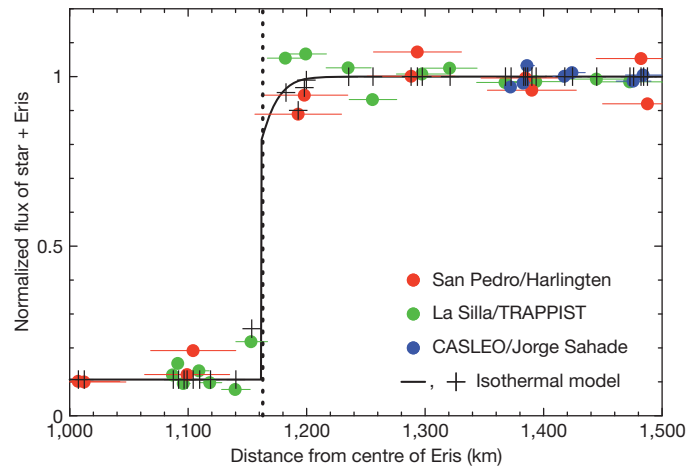
$R_E = 1,500$  km previously found<sup>11</sup> can be ruled out, as it requires fine tunings in both Eris' limb and pole orientations (Fig. 2). Thus, the most straightforward interpretation of our observations is that Eris is close to spherical, remembering that larger sizes are possible in a narrow region of the parameter space. Consequently, Eris is close in size to Pluto, whose radius<sup>3–8</sup> is estimated between 1,150 and 1,200 km.

Our radius value implies a density of  $\rho = 2.52 \pm 0.05$  g cm<sup>-3</sup>, when combined with Eris' mass<sup>12</sup>. This is comparable to Haumea's density<sup>16,17</sup> ( $\sim 2.6$  g cm<sup>-3</sup>), for which a typical rock/ice ratio of 0.85/0.15 is derived<sup>18</sup>. This suggests that Haumea (and thus also Eris) is a large rocky body with a thin overlying ice shell. Note that the densities of trans-Neptunian objects (TNOs) span a large range, with  $\rho$  values of 1.0, 1.6 and 2.0 g cm<sup>-3</sup> for Varuna<sup>17</sup>, Charon<sup>19</sup> and Pluto<sup>19</sup>, respectively, pointing to diverse origins and/or evolutions. Our radius value provides a geometric albedo of  $p_V = 0.96^{+0.09}_{-0.04}$  in the visible range (Supplementary Information). This makes Eris almost as bright as a perfect isotropic Lambert surface (for which  $p_V = 1$  by definition), and one of the intrinsically brightest objects of the Solar System. For comparison, Saturn's satellite Enceladus is even brighter, with a geometric albedo of  $p_V \approx 1.4$ , associated with its geologically active surface<sup>20</sup>. In contrast, Eris' brightness and lack of light-curve variations may stem from the collapse of a nitrogen atmosphere (see below). We find that Eris is brighter than the TNO 2002 TX<sub>300</sub>, whose high albedo ( $0.88^{+0.15}_{-0.06}$ ) is probably due to the exposure of fresh water-ice<sup>21</sup>.

We now reassess Eris' surface temperature in the light of our new results. Measurements by the Spitzer<sup>22</sup> and IRAM<sup>11</sup> satellites imply disk-averaged brightness temperatures of  $T_b = 30 \pm 1.2$  K and  $T_b = 38 \pm 7.5$  K at 70 and 1,200  $\mu$ m, respectively. As a completely absorbing surface at Eris' distance has a temperature  $T_0 = 40$  K, the second value is surprisingly high (and consistent with the fact that the previously found radius<sup>11</sup> of 1,500 km is about 30% higher than our value), but we note that a unique brightness temperature  $T_b \approx 31$  K matches both the Spitzer (at 70  $\mu$ m) and IRAM (1,200  $\mu$ m) measurements at the  $1.5\sigma$  level (Supplementary Fig. 5). However, Eris' surface temperature is probably not uniform, because an atmosphere (if any) would be too tenuous to isothermize the surface frosts, as occurs for Triton and Pluto. We therefore consider two extreme standard temperature distribution models, corresponding to (1) a warmer slow rotator (or equivalently, pole-on orientation, or zero thermal inertia, the standard thermal model, STM) with sub-solar temperature  $T_{ss}$ , and (2) a cooler fast rotator with equator-on geometry (isothermal with latitude model, ILM), with equatorial temperature  $T_{eq}$ .

In the STM, both Spitzer and IRAM fluxes are reproduced satisfactorily with  $T_{ss} \approx 35$  K (Supplementary Fig. 5, Supplementary Tables 4 and 5). The thermal equilibrium equation  $T_{ss} = T_0[(1 - p_V q)/(\epsilon \eta)]^{1/4}$  then provides a relationship between the beaming factor  $\eta$  (describing the effects of surface roughness), the phase integral  $q$  and the surface emissivity  $\epsilon$ , where  $A = p_V q$  is the Bond albedo, which measures the fraction of reflected solar energy. Using a standard value<sup>22</sup>  $\epsilon = 0.9$  and a plausible range from  $\eta = 1$  (no roughness) to 0.7 (large surface roughness), we obtain  $q = 0.49$ – $0.66$ , fully consistent with the values for Saturn's brightest icy satellites<sup>20,23</sup>. The ILM in contrast leads to the extreme condition  $0 < q < 0.24$ , which is an implausible range as bright objects also have large phase integrals<sup>24</sup>. Essentially, the fast rotator model does not provide enough thermal flux given the new, smaller size of Eris. We therefore strongly favour the STM, implying either a pole-on orientation or a very small thermal inertia, as observed in other TNOs<sup>25,26</sup>.

The occultation puts an upper limit on a putative atmosphere around Eris. As discussed in Supplementary Information, our preferred model is an isothermal N<sub>2</sub> atmosphere near 30 K, for which we can place an upper limit of about 1 nbar ( $1\sigma$  level) at the surface (Fig. 3). Similar limits are obtained for hypothetical CH<sub>4</sub> or Ar atmospheres. Also discussed in Supplementary Information is the possibility that a Pluto-like atmosphere sublimates as Eris approaches its perihelion, at 37.8 AU from the Sun. In that case, Eris would currently be a



**Figure 3 | Upper limit on Eris' atmosphere.** Each data point (filled coloured circles) obtained at three of the stations shown in Fig. 2 has been projected onto a radial scale (distance from Eris' centre), using the circular solution 2 displayed in Fig. 2. The horizontal bars indicate the finite radial resolution associated with finite integration intervals; the vertical dotted line shows the adopted Eris radius,  $R_E = 1,163$  km. The black solid line is a model light curve obtained at 1-km resolution, using an isothermal pure N<sub>2</sub> atmosphere. Black crosses mark the expected flux associated with each data point, once the convolution with the finite integration intervals has been performed. The fit minimizes the differences between the black crosses and the corresponding data points (filled circles). The model shown here is the  $3\sigma$ -level upper limit of an isothermal N<sub>2</sub> atmospheric profile, with  $T = 27.7$  K and a surface pressure of 2.9 nbar. Most of the constraint on the model comes from the two data points obtained at La Silla (the two green filled circles just right of the vertical dotted line), corresponding to the data points obtained just before and just after the event (Fig. 1). The two closest San Pedro data points (red) have only a small contribution to the  $\chi^2$  value, while the CASLEO data points (blue) are too far away from Eris' edge ( $\sim 200$  km) to effectively constrain the atmospheric pressure. Using solution 1 instead of solution 2 for Eris' shape would have a minimal impact on the atmospheric upper limit, as this would slightly displace the San Pedro data points in the plot, leaving the La Silla points where they are shown here.

dormant Pluto twin, with a bright icy surface created by a collapsed atmosphere. Detailed models are required, however, to confirm this model.

Received 5 June; accepted 6 September 2011.

1. Tegler, S. C. *et al.* Methane and nitrogen abundances on Pluto and Eris. *Astrophys. J.* **725**, 1296–1305 (2010).
2. Yelle, R. V. & Elliot, J. L. in *Pluto and Charon* (eds Stern, S. A. & Tholen, D. J.) 347–390 (Univ. Arizona Press, 1997).
3. Elliot, J. L. *et al.* Changes in Pluto's atmosphere: 1988–2006. *Astron. J.* **134**, 1–13 (2007).
4. Lellouch, E. *et al.* Pluto's lower atmosphere structure and methane abundance from high-resolution spectroscopy and stellar occultations. *Astron. Astrophys.* **495**, L17–L21 (2009).
5. Zalucha, M. A. *et al.* An analysis of Pluto occultation light curves using an atmospheric radiative-convective model. *Icarus* **211**, 804–818 (2011).
6. Tholen, D. J. & Buie, M. W. Further analysis of Pluto-Charon mutual event observations. *Bull. Am. Astron. Soc.* **22**, 1129 (1990).
7. Young, E. F. & Binzel, R. P. A new determination of radii and limb parameters for Pluto and Charon from mutual event lightcurves. *Icarus* **108**, 186–199 (1994).
8. Tholen, D. J. & Buie, M. W. in *Pluto and Charon* (eds Stern, S. A. & Tholen, D. J.) 193–219 (Univ. Arizona Press, 1997).
9. Brown, M. E., Trujillo, C. A. & Rabinowitz, D. L. Discovery of a planetary-sized object in the scattered Kuiper Belt. *Astrophys. J.* **635**, L97–L100 (2005).
10. Brown, M. E., Schaller, E. L., Roe, H. G., Rabinowitz, D. L. & Trujillo, C. A. Direct measurement of the size of 2003 UB313 from the Hubble Space Telescope. *Astrophys. J.* **643**, L61–L63 (2006).
11. Bertoldi, F. F., Altenhoff, W., Weiss, A., Menten, K. M. & Thum, C. The trans-neptunian object UB313 is larger than Pluto. *Nature* **439**, 563–564 (2006).
12. Brown, M. E. & Schaller, E. L. The mass of dwarf planet Eris. *Science* **316**, 1585 (2007).
13. Sheppard, S. S. *et al.* Light curves of dwarf plutonian planets and other large Kuiper Belt Objects: their rotations, phase functions, and absolute magnitudes. *Astron. J.* **134**, 787–798 (2007).
14. Duffard, R. *et al.* A study of photometric variations on the dwarf planet (136199) Eris. *Astron. Astrophys.* **479**, 877–881 (2008).

15. Assafin, M. *et al.* Precise predictions of stellar occultations by Pluto, Charon, Nix and Hydra for 2008–2015. *Astron. Astrophys.* **515**, A32 (2010).
16. Rabinowitz, M. E. *et al.* Photometric observations constraining the size, shape, and albedo of 2003 EL61, a rapidly rotating, Pluto-sized object in the Kuiper Belt. *Astrophys. J.* **639**, 1238–1251 (2006).
17. Lacerda, P. & Jewitt, D. C. Densities of solar system objects from their rotational light curves. *Astron. J.* **133**, 1393–1408 (2007).
18. McKinnon, W. B., Prialnik, D., Stern, S. A. & Coradini, A. in *The Solar System Beyond Neptune* (eds Barucci, M. A., Boehnhardt, H., Cruikshank, D. P. & Morbidelli, A.) 213–241 (Univ. Arizona Press, 2008).
19. Tholen, D. J., Buie, M. W., Grundy, W. M. & Elliott, G. T. Masses of Nix and Hydra. *Astron. J.* **135**, 777–784 (2008).
20. Verisicer, A., French, R., Showalter, M. & Helfenstein, P. Enceladus: cosmic graffiti artist caught in the act. *Science* **315**, 815 (2007).
21. Elliot, J. L. *et al.* Size and albedo of Kuiper belt object 55636 from a stellar occultation. *Nature* **465**, 897–900 (2010).
22. Stansberry, J. *et al.* in *The Solar System beyond Neptune* (eds Barucci, M. A., Boehnhardt, H., Cruikshank, D. P. & Morbidelli, A.) 161–179 (Univ. Arizona Press, 2008).
23. Howett, C. J. A., Spencer, J. R., Pearl, J. & Segura, M. Thermal inertia and bolometric Bond albedo values for Mimas, Enceladus, Tethys, Dione, Rhea and Iapetus as derived from Cassini/CIRS measurements. *Icarus* **206**, 573–593 (2010).
24. Brucker, M. J. *et al.* High albedos of low inclination classical Kuiper belt objects. *Icarus* **201**, 284–294 (2009).
25. Müller, T. G. *et al.* “TNOs are cool”: a survey of the trans-Neptunian region. I. Results from the Herschel science demonstration phase (SDP). *Astron. Astrophys.* **518**, L146 (2010).
26. Lellouch, E. *et al.* “TNOs are cool”: a survey of the trans-Neptunian region. II. The thermal lightcurve of (136108) Haumea. *Astron. Astrophys.* **518**, L147 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank W. M. Owen and S. Preston for providing astrometric updates on the occulted star just before the observations, and I. Belskaya and M. E. Brown for discussions when writing the paper. We acknowledge support from the French grant ‘Beyond Neptune’ and from the Institut Universitaire de France. J.L.O., A.J.C.-T., L.C. and M.T.E. acknowledge funding from Spanish AYA grants and FEDER funds. TRAPPIST is a project funded by the Belgian Fund for Scientific Research (FRS-FNRS) with the participation of the Swiss National Science Foundation (SNF). J.I.B.C. acknowledges CNPq and FAPERJ grants. F.B.-R. acknowledges the support of CDFB/CAPES, Brazil. W.J.B.C., W.R. and F.P.S. thank the Brazilian Agency FAPEMIG.

**Author Contributions** B.S. helped plan the campaign, centralized the stellar occultation predictions, participated in the observations, analysed data, wrote and ran the diffraction, limb-fitting and ray-tracing codes, and wrote part of the paper. J.L.O. helped plan the campaign, analysed data for the prediction, participated in the observations, obtained and analysed data, and wrote part of the paper. E.L. analysed the implications of the results for the Eris thermal model, albedo constraints and putative atmospheric structure, and wrote part of the paper. M.A., F.B.-R., A.H.A., J.I.B.C., R.V.M., D.N.d.S.N. and R.B. discovered the star candidate and analysed data for the predictions. E.J. and A.M. obtained and analysed the positive occultation detection at La Silla/TRAPPIST and San Pedro/Harlingen telescopes, respectively. F.B.-R., F.C., M.G. and J.M. analysed data, D.H. calculated Dysnomia’s position at the moment of occultation and wrote part of the paper. All other authors participated in the planning of the campaign and/or the observations, and the authors listed in Supplementary Table 2 were responsible for the observations. All authors were given the opportunity to review the results and comment on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to B.S. ([bruno.sicardy@obspm.fr](mailto:bruno.sicardy@obspm.fr)).

# Laser cooling and real-time measurement of the nuclear spin environment of a solid-state qubit

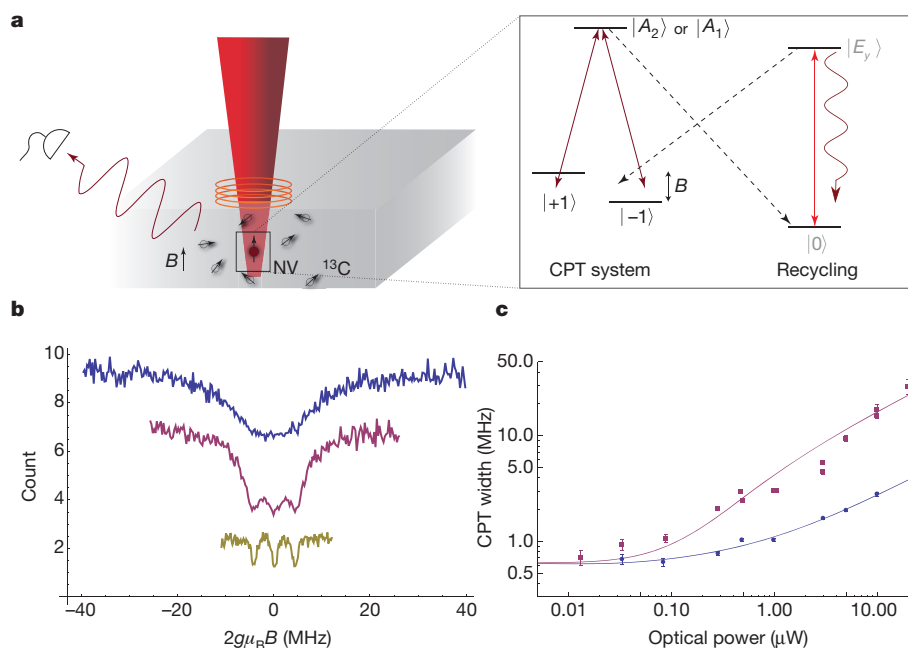
E. Togan<sup>1</sup>\*, Y. Chu<sup>1</sup>\*, A. Imamoglu<sup>2</sup> & M. D. Lukin<sup>1</sup>

Control over quantum dynamics of open systems is one of the central challenges in quantum science and engineering. Coherent optical techniques, such as coherent population trapping involving dark resonances<sup>1,2</sup>, are widely used to control quantum states of isolated atoms and ions. In conjunction with spontaneous emission, they allow for laser cooling of atomic motion<sup>3</sup>, preparation and manipulation of atomic states<sup>4</sup>, and rapid quantum optical measurements that are essential for applications in metrology<sup>5–7</sup>. Here we show that these techniques can be applied to monitor and control individual atom-like impurities, and their local environment<sup>8–11</sup>, in the solid state. Using all-optical manipulation of the electronic spin of an individual nitrogen–vacancy colour centre in diamond, we demonstrate optical cooling, real-time measurement and conditional preparation of its nuclear spin environment by post-selection. These methods offer potential applications ranging from all-optical nanomagnetometry to quantum feedback control of solid-state qubits, and may lead to new approaches for quantum information storage and processing

Over the past two decades, coherent population trapping (CPT) has been used in the laser cooling of neutral atoms and ions<sup>3</sup>, the creation

of ultracold molecules<sup>4</sup>, optical magnetometry<sup>5,6</sup>, and atomic clocks<sup>7</sup>, as well as in slowing and stopping light pulses<sup>2</sup>. The electronic spin of the nitrogen–vacancy (NV) centre is a promising system for extending these techniques to the solid state. The NV centre has a long-lived spin triplet as its electronic ground state<sup>12</sup>, whose  $m_S = \pm 1$  and  $m_S = 0$  sublevels we denote by  $|\pm 1\rangle$  and  $|0\rangle$ . In pure samples, the electron spin dynamics are governed by interactions with the spin-1  $^{14}\text{N}$  nucleus of the NV centre and spin-1/2  $^{13}\text{C}$  nuclei present in 1.1% natural abundance in the diamond lattice (Fig. 1a). Control over nuclear spins<sup>13,14</sup> is of interest both for fundamental studies and in applications such as nanoscale magnetic sensing<sup>15,16</sup> and the realization of quantum networks<sup>17,18</sup>. Here we achieve such control by two complementary methods: effective cooling of nuclear spins through nuclear-state-selective CPT<sup>8</sup>, and conditional preparation based on fast measurements of the nuclear environment and subsequent post-selection<sup>11</sup>.

Whereas most previous work involved the use of microwave and radio-frequency fields to manipulate both the electronic and the nuclear spin states, we use all-optical control of the electronic spin<sup>19–21</sup>. Specifically, we make use of  $\Lambda$ -type level configurations involving the NV centre's optically excited electronic states  $|A_1\rangle$  and  $|A_2\rangle$  and the

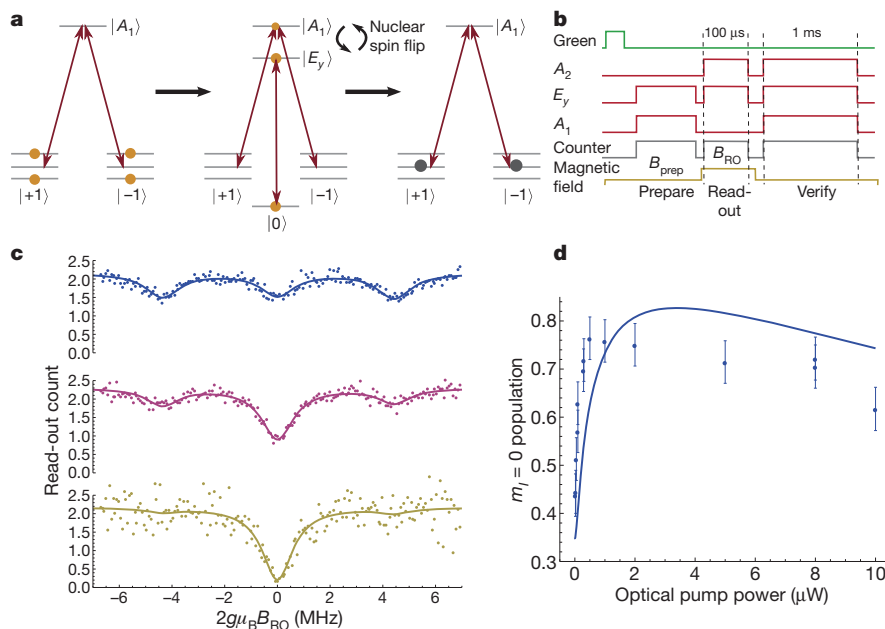


**Figure 1 | Coherent population trapping in NV centres.** **a**, The  $\Lambda$ -type transitions between the ground states  $|\pm 1\rangle$  and excited states  $|A_1\rangle$  and  $|A_2\rangle$  of a single NV centre are addressed with a CPT laser, while a recycling laser drives the  $|0\rangle \rightarrow |E_y\rangle$  transition. An external magnetic field,  $B$ , is applied using a solenoid. **b**, Photon count from NV centre NVa in a 300- $\mu\text{s}$  window are plotted versus the applied field for three different powers of a laser addressing the state

$|A_2\rangle$ : blue, 10  $\mu\text{W}$ ; pink, 3  $\mu\text{W}$ ; yellow, 0.1  $\mu\text{W}$ . The blue and pink data sets are shifted vertically by five and, respectively, two counts for clarity.  $\mu_B$ , Bohr magneton;  $g$ , gyromagnetic ratio. **c**, Width of individual  $^{14}\text{N}$  CPT lines versus CPT laser power when the  $|A_1\rangle$  (blue) or  $|A_2\rangle$  (pink) state is used. Solid curves represent the theoretical model discussed in the main text and Supplementary Information. Error bars, s.d.

<sup>1</sup>Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>2</sup>Institute for Quantum Electronics, ETH-Zurich, CH-8093 Zurich, Switzerland.

\*These authors contributed equally to this work.



**Figure 2 | Optical control and conditional preparation of the proximal  $^{14}\text{N}$  nuclear spin.** **a**, Mechanism for optical pumping of  $^{14}\text{N}$  states. **b**, Pulse sequence for  $^{14}\text{N}$  optical pumping using the laser addressing the state  $|A_1\rangle$  ( $A_1$  laser) and a fixed read-out of the prepared state using the  $A_2$  laser. To ensure that the nitrogen–vacancy was not ionized for all subsequent data runs, we turn on all three lasers at the end of each run so that there is no dark state, and only keep data from runs in which we obtain a high number of counts during this verification step. **c**, Counts collected with NVa during the read-out step versus the read-out magnetic field when there is no preparation step (blue) and when

there is preparation with optical pumping using 100 nW of  $A_1$  laser power for 1.9 ms (pink). The yellow curve shows the results of  $^{14}\text{N}$  polarization through measurement-based preparation by selecting the read-out events in which the number of counts collected during the last 500  $\mu\text{s}$  of preparation is zero (Supplementary Information, section 4.2). **d**, Steady-state population in the  $m_I = 0$  state after optical pumping for varying powers of the  $A_1$  laser, with theoretical model described in Supplementary Information (solid line). Error bars, s.d.

$|\pm 1\rangle$  ground states<sup>18,22</sup> (Fig. 1a). At low temperatures ( $<10\text{ K}$ ) and in the limit of zero strain,  $|A_1\rangle$  and  $|A_2\rangle$  are entangled states of spin and orbital momentum, both coupled to  $|+1\rangle$  with  $\sigma_-$ -polarized light and to  $|-1\rangle$  with  $\sigma_+$ -polarized light. Correspondingly, excitation with linearly polarized light drives the NV centre into a ‘dark’ superposition state when the two-photon detuning is zero<sup>1</sup>. In the present case, the two-photon detuning is determined by the Zeeman splitting between the  $|\pm 1\rangle$  states due to the combined effect of the Overhauser field originating in the nuclear spin environment and any externally applied magnetic field<sup>8,10</sup>. When the external field exactly compensates the Overhauser field, the electronic spin of the NV centre is pumped into the dark state after a few optical cycles and remains in there, resulting in vanishing fluorescence. This is the basis of the dark resonances and CPT.

In our experiments,  $|A_1\rangle$  and  $|A_2\rangle$  are separated in energy by an amount corresponding to  $\sim 3\text{ GHz}$  and are addressed individually with a single, linearly polarized laser at near-zero magnetic field. Because there is a finite branching ratio from the  $m_S = \pm 1$  manifold of the electronic spin into state  $|0\rangle$ , we use a recycling laser that drives the transition between  $|0\rangle$  and the excited state  $|E_g\rangle$ , which decays with a small but non-vanishing probability ( $\sim 10^{-2}$ ) back to the states  $|\pm 1\rangle$ . In Fig. 1b, we present our observations of the CPT spectrum as a function of an external magnetic field at three different powers of a laser tuned to the  $|\pm 1\rangle \rightarrow |A_2\rangle$  transition. A broad resonance is observed at high powers, but as the power is reduced we clearly resolve three features in the spectrum, separated by 4.4 MHz, which is twice the hyperfine splitting between three  $^{14}\text{N}$  nuclear spin states. This separation corresponds to the magnetic field required to bring the electronic  $m_S = \pm 1$  hyperfine states with equal nuclear spin projections ( $m_I = \pm 1, 0$ ) into two-photon resonance.

The dependence of the CPT resonance width on the laser power (Fig. 1c) shows that repumping on the transition  $|0\rangle \leftrightarrow |E_g\rangle$  has an important role in our experiment. By contrast with a conventional, closed three-level system, this recycling transition can be used to enhance

the utility of our CPT system by both decreasing the width of the CPT resonance and increasing the signal-to-noise ratio. The state  $|A_1\rangle$  decays into the  $m_S = 0$  ground state through the NV centre’s metastable singlet state with a substantial probability, of  $\sim 40\%$  (Supplementary Information). However, the population returns to the  $m_S = \pm 1$  state from  $|E_g\rangle$  after only  $\sim 100$  optical excitation cycles. As a result, away from the two-photon resonance, the NV centre quickly decays to  $|0\rangle$  after being excited, and then scatters many photons through the transition  $|0\rangle \leftrightarrow |E_g\rangle$  before returning to the  $\Lambda$  system (the  $\Lambda$ -type level configuration described above). If the NV centre is not in a dark state, this process effectively increases the number of photons we collect by a factor of  $2/\eta = \gamma_{sl}/\gamma_{ce}$ , where  $\gamma_{ce}$  is the cross-transition rate from  $|E_g\rangle$  into  $|\pm 1\rangle$  and  $\gamma_{sl}$  is the transition rate from  $|A_1\rangle$  to the singlet. The cycling effect also reduces the width of the CPT line because the  $|0\rangle \leftrightarrow |E_g\rangle$  transition quickly saturates away from two-photon resonance, provided that the CPT laser excitation rate exceeds the rate of leakage out of the recycling transition. Significantly, both of these effects lead to improved sensitivity of dark resonances to small changes in two-photon detuning.

To demonstrate this increase in sensitivity, in Fig. 1c we compare the widths of dark resonances observed through excitation of  $|A_1\rangle$  and  $|A_2\rangle$ . Through independent measurements of the branching ratios (Supplementary Information), we determined that  $|A_1\rangle$  and  $|A_2\rangle$  correspond to an open and, respectively, nearly closed  $\Lambda$  systems, with  $\eta_{A_1} \approx 3.1 \times 10^{-2}$  and  $\eta_{A_2} \approx 2.6$ . We compare these experimental results with a theoretical model described in Supplementary Information, which predicts that the resonance linewidth,  $\delta_0$ , is given by

$$\delta_0 = \sqrt{\frac{R_A^2}{1 + (1/\eta)(R_A/R_E + 2R_A/\gamma)}} \approx \sqrt{R_A R_E \eta \gamma / (R_E + \gamma)}$$

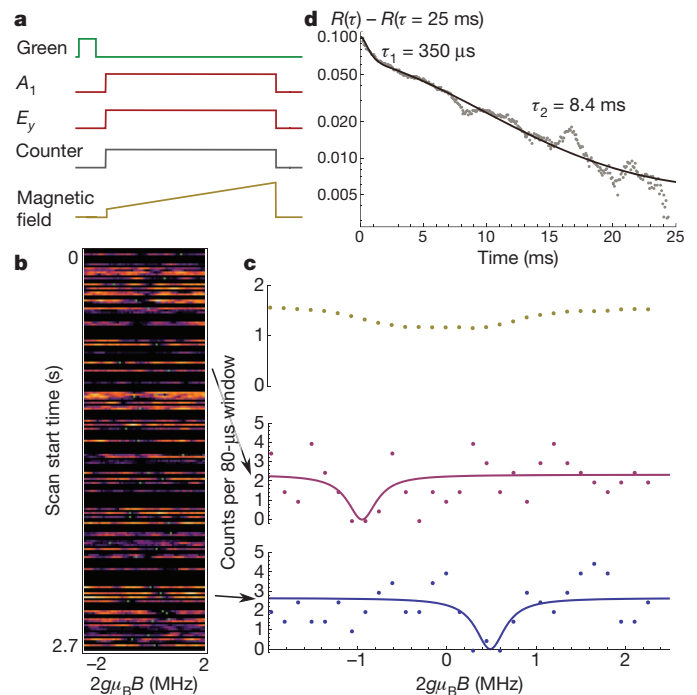
for small  $\eta$ , where  $R_A$  and  $R_E$  correspond to the optical excitation rates of a laser tuned to states  $|A_1\rangle$  or  $|A_2\rangle$  and, respectively,  $|E_y\rangle$ , and  $\gamma$  is the decay rate of excited states. The width at low powers is determined by the random magnetic field associated with surrounding  $^{13}\text{C}$  nuclear states. When this line broadening mechanism is taken into account (Supplementary Information), the experimental results are in excellent agreement with these predictions, which we plot as solid lines in Fig. 1c.

Having resolved the hyperfine coupling between the nitrogen-vacancy electron spin and  $^{14}\text{N}$  spin, we now demonstrate optical cooling of the nuclear spin states using dark resonances. This method (Fig. 2a) is reminiscent of laser cooling of atomic motion by velocity-selective CPT<sup>3,8</sup>. A redistribution of the  $^{14}\text{N}$  spin state population on optical excitation takes place because the hyperfine coupling in the excited electronic state of the NV centre is enhanced by a factor of  $\sim 20$  relative to that in the ground state<sup>23</sup>. If the external field is set such that, for example, the  $m_I = 0$  hyperfine states are in two-photon resonance, only the states with nuclear configuration  $m_I = \pm 1$  will be promoted to the excited states, where flip-flops with the electron spin will change the nuclear spin state to  $m_I = 0$ . When the NV centre spontaneously decays into the dark superposition of electronic spin states, optical excitation will cease, resulting in effective polarization (cooling) of nuclear spin into the  $m_I = 0$  state.

In Fig. 2c, we present our observations of laser cooling of  $^{14}\text{N}$  nuclear spins by CPT. For each point, the pulse sequence shown in Fig. 2b is applied, where the magnetic field  $B_{\text{prep}}$  is kept at zero during the preparation/optical pumping process, and fluorescence is collected when the field is switched to a particular read-out value,  $B_{\text{RO}}$ . The increased contrast of the  $m_I = 0$  CPT line relative to the other two corresponds to a nuclear spin polarization of  $61.5 \pm 4.4\%$ . As shown in Fig. 2d, by optimizing the power of the laser tuned to  $|A_1\rangle$  we achieve a maximum nuclear polarization of  $76.4 \pm 4.4\%$  over a timescale of  $353 \pm 34 \mu\text{s}$ . The degree of polarization is probably limited by the escape rate out of the dark state due to off-resonance excitation of  $|A_2\rangle$  and dephasing caused by surrounding  $^{13}\text{C}$  nuclei. A simple theoretical model taking into account these two processes and using independently measured parameters (Fig. 2d, solid line) reproduces the qualitative features of our experimental results.

We can further improve the preparation of  $^{14}\text{N}$  nuclear spins in a desired state by measurement and post-selection, as predicted by theoretical proposals<sup>10,11,24</sup>. Specifically, the observation of zero photo-detection events during the preparation step at  $B_{\text{prep}} = 0$  determines the  $^{14}\text{N}$  spins to be in the  $m_I = 0$  state. For instances in which such a measurement result was obtained, the bottom plot of Fig. 2c shows the nuclear spin populations measured during the subsequent probe step. The resulting  $^{14}\text{N}$  polarization, of more than  $92 \pm 6\%$ , demonstrates that high-fidelity conditional preparation of nuclear spins is possible.

We next extend our technique to monitor and cool the many-body environment of the NV centre, which consists of  $^{13}\text{C}$  nuclei distributed throughout the diamond lattice. The large number of nuclear spin configurations associated with an unpolarized environment results in a random Overhauser field,  $B_{\text{Ov}}$ , with unresolved hyperfine lines. This produces a finite CPT linewidth in measurements that average over all configurations of the  $^{13}\text{C}$  spin bath (Fig. 1b, c). We can overcome this limitation by making fast measurements. The key idea of our approach is to use the long correlation time,  $T_1^{\text{nuc}}$ , associated with evolution of the nuclear bath, to observe its instantaneous state and its dynamics. Such a fast measurement is illustrated in Fig. 3a, where the externally applied field is ramped across a single  $^{14}\text{N}$   $m_I = 0$  line while the CPT lasers are on. The yellow points in Fig. 3c show the line shape averaged over many experimental runs, and the intensity plot in Fig. 3b shows counts collected in  $80\text{-}\mu\text{s}$  time bins during successive individual runs, many of which distinctly show a narrow dark region. Lorentzian fits to selected experimental scans (Fig. 3c, blue and pink curves) reveal ‘instantaneous’ CPT resonances with linewidths that are more than a factor three less than those of the averaged measurement.

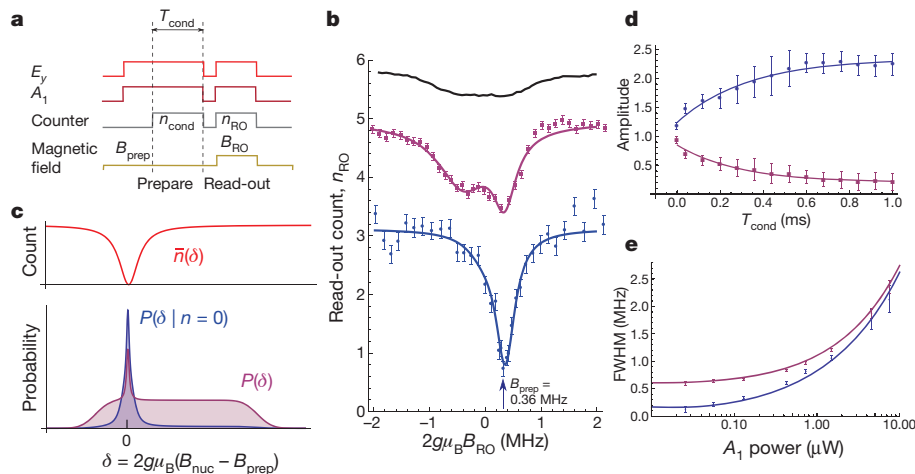


**Figure 3 | Observation of instantaneous Overhauser field from the  $^{13}\text{C}$  spin bath.** **a**, Pulse sequence for real-time measurement of the  $^{13}\text{C}$  nuclear configuration. The applied magnetic field is ramped over a single  $^{14}\text{N}$  CPT line over 5 ms while counts are collected in  $80\text{-}\mu\text{s}$  bins. **b**, Counts from 200 successive runs are shown on horizontal lines for NVb. Runs in which the verification step fails are blacked out. The centres of constrained Lorentzian fits (Supplementary Information, section 7) to individual runs are indicated with green dots. **c**, Two such individual runs are shown with their fits (pink and blue), along with an average of scans that passed verification (yellow). **d**, Autocorrelation,  $R(\tau)$ , of counts with magnetic field fixed at the  $m_I = 1$   $^{14}\text{N}$  line. The fit is to a bi-exponential decay.

The change in the centres of the dark regions (Fig. 3b, green dots) indicates that the instantaneous field evolves in time.

To provide more quantitative insight into the dynamics of the nuclear environment, we record the fluorescence counts at a fixed value of the external magnetic field with a time resolution of  $80\mu\text{s}$  during 50-ms time intervals. The autocorrelation of the resulting photon detection events (Fig. 3d) reveals two distinct timescales:  $\tau_1 = 350 \pm 30 \mu\text{s}$ , consistent with the  $^{14}\text{N}$  nuclear spin polarization timescale, and  $\tau_2 = 8.40 \pm 0.20 \text{ ms}$ . Most notably, because we can detect dark states of the NV centre with a resolution of  $80 \mu\text{s}$ , these results indicate that reliable measurement of the Overhauser field is possible within its correlation time.

We now demonstrate how fast measurements can be used to prepare conditionally the  $^{13}\text{C}$  environment of the NV centre in a desired state with post-selection. We record counts accumulated during both the preparation step and the read-out step with relatively low power using the sequence shown in Fig. 4a. Similar to measurement-based preparation of the  $^{14}\text{N}$  spin, by conditionally selecting zero-photon-detection events during the preparation step, we can select the states of the  $^{13}\text{C}$  environment that have vanishing two-photon detuning ( $\delta = 2g\mu_B(B_{\text{prep}} + B_{\text{Ov}}) = 0$ ). The pink curve in Fig. 4b shows (unconditioned) read-out counts recorded following the preparation step, whereas the blue curve shows the results of (conditioned) measurement-based preparation. The measured width of such a conditionally prepared distribution is significantly smaller than the width corresponding to individual  $^{14}\text{N}$  resonances obtained without preparation. We find that although this width depends on  $B_{\text{prep}}$ , so too does the position of the narrow feature, indicating that we can conditionally prepare the  $^{13}\text{C}$  environment by post-selection with a



**Figure 4 | Measurement-based preparation of  $^{13}\text{C}$  spin bath.** **a**, Pulse sequence for preparation and subsequent measurement of the  $^{13}\text{C}$  configuration.  $B_{\text{prep}}$  is set within the central  $^{14}\text{N}$  line and  $B_{\text{RO}}$  is varied to cover all associated  $^{13}\text{C}$  states. A green laser pulse and a  $^{14}\text{N}$  optical pumping step with the  $A_2$  and  $E_y$  lasers (not shown) occur before this. Counts during a conditioning window of length  $T_{\text{cond}}$  ( $n_{\text{cond}}$ ) at the end of preparation and during the 500- $\mu\text{s}$  read-out window ( $n_{\text{RO}}$ ) are recorded for each run. The data presented are an average of many such experimental runs. **b**,  $n_{\text{RO}}$  for NVb versus  $B_{\text{RO}}$  is shown in pink with double-Lorentzian fit. The same data set analysed by keeping only events with  $n_{\text{cond}} = 0$  is shown in blue. The

unprepared  $^{13}\text{C}$  distribution (the same as yellow curve in Fig. 3c) is shown in black for qualitative comparison (shifted by 4.3 counts for clarity). **c**, Physics of conditional preparation through measurement. **d**, Amplitudes of broad (blue) and narrow (pink) distributions versus  $T_{\text{cond}}$  for NVb. The same data set was used for each point and the length of the conditioning window was changed in post-processing. **e**, Full-width at half-maximum (FWHM) of measured  $^{13}\text{C}$  distribution with (blue) and without (pink) conditional preparation, versus  $A_1$  laser power for NVa. Solid lines are theoretical predictions for the read-out linewidth,  $\delta_0$  (blue), and the CPT linewidth for the unprepared  $^{13}\text{C}$  distribution (pink; same as in Fig. 1c). Error bars, s.d.

Overhauser field of our choice (Supplementary Fig. 8). The prepared configurations are long lived both when unilluminated ( $\gg 6$  ms; Supplementary Fig. 9) and in the presence of laser light, consistent with autocorrelation measurements ( $\tau_2 = 8.4$  ms; Fig. 3d).

We now discuss our experimental results and explore the limits of our ability to probe and prepare the  $^{13}\text{C}$  environment. As discussed in Methods Summary and illustrated in Fig. 4c, conditional measurement prepares an Overhauser field distribution that consists of the broad unconditioned distribution suppressed by  $\exp(-CT_{\text{cond}})$  and a narrow peak with a width  $\delta_c = \sqrt{\ln(2)\delta_0/\sqrt{CT_{\text{cond}}}}$ , where  $C$  is the fluorescence rate of the bright state and  $T_{\text{cond}}$  is the measurement time. The read-out step itself has a ‘resolution’ determined by the dark-resonance linewidth,  $\delta_0$ . The observed features represent a convolution of the dark-resonance probe with the conditionally prepared distribution. For the conditional preparation to be effective, we require that  $CT_{\text{cond}} > 1$  and, therefore,  $\delta_c < \delta_0$ , indicating that the measured CPT linewidth will be limited by the read-out step. Experimentally, we find that our measured line shapes can be well fitted by a combination of two Lorentzian distributions, one narrow and one broad, whose widths and positions are only weakly dependent on photodetection time,  $T_{\text{cond}}$ . However, as  $T_{\text{cond}}$  is increased, the relative weight of the narrow distribution increases (Fig. 4d). This is consistent with the theoretical prediction that the read-out-limited width of the narrow resonance does not depend on  $T_{\text{cond}}$ , and better discrimination in conditional measurements increases the probability that the nuclear spin state is prepared in the narrow distribution.

Notably, we find that even without conditioning (Fig. 4b, pink line), a narrow distribution of nuclear spin configurations around  $B_{\text{prep}}$  is prepared. This modification of the nuclear distribution is a result of CPT-based laser cooling of the  $^{13}\text{C}$  bath, consistent with the predictions of ref. 8. The specific physical mechanism of such cooling probably involves electronic-spin-dependent evolution of the  $^{13}\text{C}$  nuclei, and will be discussed in detail in future studies. We emphasize that this observation provides a clear indication that the magnetic environment is affected by the dynamics of the NV centre, providing direct evidence that the nitrogen-vacancy spin dynamics is dominated by the Overhauser field rather than external magnetic field fluctuations.

Figure 4e shows how the observed linewidth of the narrow feature depends on the CPT laser power. At low powers, the observed width reaches a minimum value of  $104 \pm 49$  kHz. This limiting width results from the effects of strain splitting of states  $|\pm 1\rangle$  on the read-out process at zero magnetic field<sup>25</sup> (see Supplementary Information for a quantitative discussion of effects of strain). Owing to this splitting, very small magnetic field changes do not shift the energies of states  $|\pm 1\rangle$  to first order. Therefore, our CPT read-out signal becomes insensitive to Zeeman shifts of the order of twice the strain splitting (Supplementary Information). In addition, a minimal linewidth of  $\sim 400$  kHz was obtained for measurements performed with a separate NV centre (NVb) subject to higher strain.

The limit associated with strain splitting can be easily circumvented by using a large external magnetic field to split the spin states  $|\pm 1\rangle$  and two laser frequencies to address the NV centre in a Raman configuration near zero two-photon detuning. As described in Supplementary Information, a modest increase in collection efficiency by, say, one order of magnitude<sup>26</sup> would allow us to obtain quantum-limited narrowing of the nuclear distribution to  $\delta_c \approx 1/T_{\text{cond}}$ , which, in turn, could be on the order of the inverse lifetime of the given nuclear configuration<sup>8</sup>.

The experiments reported here offer the intriguing prospect of using coherent optical techniques to control nuclear spins surrounding quantum emitters. For instance, the technique that we describe can be used to study the quantum many-body dynamics of ‘central-spin’ models in real time, either in isolation or in the presence of dissipation<sup>27</sup>. Specific examples of this include nuclear field diffusion that, in the presence of CPT lasers, is expected to have statistical properties reminiscent of Lévy flights in velocity-selective CPT<sup>28</sup>. Furthermore, our approach allows for direct application of quantum feedback control to drive nuclear spins deterministically into a desired state. This may be used to prepare non-classical superposition states of nuclear spins analogous to spin-squeezed states in atomic ensembles<sup>29</sup>, and to ‘engineer’ collective dissipation in nuclear spin ensembles useful for applications in quantum information science, such as the long-term storage of quantum states<sup>30</sup>. Finally, our method allows for an all-optical approach to magnetic sensing<sup>5,6</sup> that may have interesting applications in nanoscience<sup>15,16</sup>.

## METHODS SUMMARY

**Sample description.** The diamond sample used was a natural, high-purity, type-IIa diamond with a (111) cut, kept at  $\sim 7$  K. The data presented in the main text and the Supplementary Information came from measurements on three separate NV centres in this sample. The first NV centre (NVa) was under relatively low strain and had a narrow distribution of  $^{13}\text{C}$  states. The second, higher-strain NV centre (NVb) had a broader distribution of  $^{13}\text{C}$  states. All experiments on optical cooling and conditional preparation were repeated using both of these NV centres, with consistent results. Figures 1, 2 and 4e and Supplementary Fig. 3 present measurements on NVa. The remaining figures in the main text and Supplementary Information, excluding Supplementary Fig. 5, present measurements on NVb. The third NV centre (NVc) was used for electron spin resonance measurements (Supplementary Information, section 5), with which we calibrated the ground-state strain for NVa and NVb by assuming that it is proportional to the strain measured using the excited states  $|E_x\rangle$  and  $|E_y\rangle$  (ref. 25).

**Measurement-based preparation of  $^{13}\text{C}$  environment.** We consider the situation in which the NV centre is continuously monitored for a time  $T_{\text{cond}}$ . The average number of photons detected during preparation is given by  $\bar{n}(\delta)T_{\text{cond}}$ , where the photon detection rate,  $\bar{n}(\delta) = C\delta^2/(\delta_0^2 + \delta^2)$ , is related to the instantaneous value of the Overhauser field through the two-photon detuning,  $\delta$  (Fig. 4c). Directly after preparation, the nuclear state probability distribution determined by conditioning on obtaining zero counts is  $P(\delta | n=0)$ , which is related to the conditional probability of a zero-count event,  $P(n=0 | \delta)$ , by  $P(\delta | n=0) = P(n=0 | \delta)P(\delta)/P(n=0)$ , where  $P(n)$  and  $P(\delta)$  are respectively the unconditional count and detuning distributions. For a Poisson-distributed random process of photon counts, we find that  $P(\delta | n=0) \approx \exp(-CT_{\text{cond}}\delta^2/(\delta_0^2 + \delta^2))P(\delta)$ . As  $T_{\text{cond}}$  increases, the range of  $\delta$  for which we obtain zero counts owing to the existence of a dark state becomes small, and, for large  $\delta$ , we expect the average number of counts to be high and the probability of detecting  $n=0$  counts due to shot noise to be small. This effectively reduces the width of the conditionally prepared nuclear spin distribution.

Received 13 June; accepted 31 August 2011.

- Scully, M. O. & Zubairy, M. S. *Quantum Optics* 222–225 (Cambridge Univ. Press, 1997).
- Fleischhauer, M., Imamoglu, A. & Marangos, J. P. Electromagnetically induced transparency: optics in coherent media. *Rev. Mod. Phys.* **77**, 633–673 (2005).
- Aspect, A., Arimondo, E., Kaiser, R., Vansteenkiste, N. & Cohen-Tannoudji, C. Laser cooling below the one-photon recoil energy by velocity-selective coherent population trapping. *Phys. Rev. Lett.* **61**, 826–829 (1988).
- Ni, K.-K. *et al.* A high phase-space-density gas of polar molecules. *Science* **322**, 231–235 (2008).
- Scully, M. O. & Fleischhauer, M. High-sensitivity magnetometer based on index-enhanced media. *Phys. Rev. Lett.* **69**, 1360–1363 (1992).
- Budker, D. & Romalis, M. Optical magnetometry. *Nature Phys.* **3**, 227–234 (2007).
- Vanier, J. Atomic clocks based on coherent population trapping: a review. *Appl. Phys. B* **81**, 421–442 (2005).
- Issler, M. *et al.* Nuclear spin cooling using Overhauser-field selective coherent population trapping. *Phys. Rev. Lett.* **105**, 267202 (2010).
- Xu, X. *et al.* Optically controlled locking of the nuclear field via coherent dark-state spectroscopy. *Nature* **459**, 1105–1109 (2009).
- Stepanenko, D., Burkard, G., Giedke, G. & Imamoglu, A. Enhancement of electron spin coherence by optical preparation of nuclear spins. *Phys. Rev. Lett.* **96**, 136401 (2006).
- Giedke, G., Taylor, J. M., D'Alessandro, D., Lukin, M. D. & Imamoglu, A. Quantum measurement of a mesoscopic spin ensemble. *Phys. Rev. A* **74**, 032316 (2006).
- Manson, N. B., Harrison, J. P. & Sellars, M. J. Nitrogen-vacancy center in diamond: model of the electronic structure and associated dynamics. *Phys. Rev. B* **74**, 104303 (2006).
- Dutt, M. V. G. *et al.* Quantum register based on individual electronic and nuclear spin qubits in diamond. *Science* **316**, 1312–1316 (2007).
- Neumann, P. *et al.* Single-shot readout of a single nuclear spin. *Science* **329**, 542–544 (2010).
- Maze, J. *et al.* Nanoscale magnetic sensing with an individual electronic spin in diamond. *Nature* **455**, 644–647 (2008).
- Balasubramanian, G. *et al.* Nanoscale imaging magnetometry with diamond spins under ambient conditions. *Nature* **455**, 648–651 (2008).
- Childress, L., Taylor, J. M., Sørensen, A. S. & Lukin, M. D. Fault-tolerant quantum communication based on solid-state photon emitters. *Phys. Rev. Lett.* **96**, 070504 (2006).
- Togan, E. *et al.* Quantum entanglement between an optical photon and a solid-state spin qubit. *Nature* **466**, 730–734 (2010).
- Robledo, L., Bernien, H., van Weperen, I. & Hanson, R. Control and coherence of the optical transition of single nitrogen vacancy centers in diamond. *Phys. Rev. Lett.* **105**, 177403 (2010).
- Santori, C. *et al.* Coherent population trapping of single spins in diamond under optical excitation. *Phys. Rev. Lett.* **97**, 247401 (2006).
- Buckley, B. B., Fuchs, G. D., Bassett, L. C. & Awschalom, D. D. Spin-light coherence for single-spin measurement and control in diamond. *Science* **330**, 1212–1215 (2010).
- Maze, J. R. *et al.* Properties of nitrogen-vacancy centers in diamond: the group theoretic approach. *N. J. Phys.* **13**, 025025 (2011).
- Fuchs, G. D. *et al.* Excited-state spectroscopy using single spin manipulation in diamond. *Phys. Rev. Lett.* **101**, 117601 (2008).
- Klauser, D., Coish, W. A. & Loss, D. Nuclear spin state narrowing via gate-controlled Rabi oscillations in a double quantum dot. *Phys. Rev. B* **73**, 205302 (2006).
- Dolde, F. *et al.* Electric-field sensing using single diamond spins. *Nature Phys.* **7**, 459–463 (2011).
- Siyushev, P. *et al.* Monolithic diamond optics for single photon detection. *Appl. Phys. Lett.* **97**, 241902 (2010).
- Chen, G., Bergman, D. L. & Balents, L. Semiclassical dynamics and long-time asymptotics of the central-spin problem in a quantum dot. *Phys. Rev. B* **76**, 045312 (2007).
- Bardou, F., Bouchaud, J.-P., Aspect, A. & Cohen-Tannoudji, C. *Lévy Statistics and Laser Cooling: How Rare Events Bring Atoms to Rest* (Cambridge Univ. Press, 2002).
- Rudner, M. S., Vandersypen, L. M. K., Vuletic, V. & Levitov, L. S. Generating entanglement and squeezed states of nuclear spins in quantum dots. Preprint at (<http://arxiv.org/abs/1101.3370>) (2011).
- Verstraete, F., Wolf, M. M. & Cirac, J. I. Quantum computation and quantum-state engineering driven by dissipation. *Nature Phys.* **5**, 633–636 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank A. Aspect, J. I. Cirac, G. Giedke, M. Gullans, J. Hodges, M. Issler, V. Jacques, F. Jelezko, E. Kessler, A. Kubanek, J. McArthur, A. Pick, A. Sipahigil, J. Taylor, S. Yelin and A. S. Zibrov for discussions and experimental help. This work was supported by the NSF, NSF-funded CUA, DARPA (QUEST and QUASAR programmes), ARO MURI, the NDSEG Fellowship, the Packard Foundation and an ERC Advanced Investigator Grant. The content of this paper does not necessarily reflect the position or the policy of the US government, and no official endorsements should be inferred.

**Author Contributions** All authors contributed extensively to the work presented in this paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.D.L. ([lukin@fas.harvard.edu](mailto:lukin@fas.harvard.edu)).

# Structure and reactivity of a mononuclear non-haem iron(III)–peroxo complex

Jaehung Cho<sup>1</sup>, Sujin Jeon<sup>1</sup>, Samuel A. Wilson<sup>2</sup>, Lei V. Liu<sup>2</sup>, Eun A. Kang<sup>1</sup>, Joseph J. Braymer<sup>3</sup>, Mi Hee Lim<sup>3</sup>, Britt Hedman<sup>4</sup>, Keith O. Hodgson<sup>2,4</sup>, Joan Selverstone Valentine<sup>1,5</sup>, Edward I. Solomon<sup>2,4</sup> & Wonwoo Nam<sup>1</sup>

Oxygen-containing mononuclear iron species—iron(III)–peroxo, iron(III)–hydroperoxo and iron(IV)–oxo—are key intermediates in the catalytic activation of dioxygen by iron-containing metallo-enzymes<sup>1–7</sup>. It has been difficult to generate synthetic analogues of these three active iron–oxygen species in identical host complexes, which is necessary to elucidate changes to the structure of the iron centre during catalysis and the factors that control their chemical reactivities with substrates. Here we report the high-resolution crystal structure of a mononuclear non-haem side-on iron(III)–peroxo complex,  $[\text{Fe}(\text{III})(\text{TMC})(\text{OO})]^+$ . We also report a series of chemical reactions in which this iron(III)–peroxo complex is cleanly converted to the iron(III)–hydroperoxo complex,  $[\text{Fe}(\text{III})(\text{TMC})(\text{OOH})]^{2+}$ , via a short-lived intermediate on protonation. This iron(III)–hydroperoxo complex then cleanly converts to the ferryl complex,  $[\text{Fe}(\text{IV})(\text{TMC})(\text{O})]^{2+}$ , via homolytic O–O bond cleavage of the iron(III)–hydroperoxo species. All three of these iron species—the three most biologically relevant iron–oxygen intermediates—have been spectroscopically characterized; we note that they have been obtained using a simple macrocyclic ligand. We have performed relative reactivity studies on these three iron species which reveal that the iron(III)–hydroperoxo complex is the most reactive of the three in the deformylation of aldehydes and that it has a similar reactivity to the iron(IV)–oxo complex in C–H bond activation of alkylaromatics. These reactivity results demonstrate that iron(III)–hydroperoxo species are viable oxidants in both nucleophilic and electrophilic reactions by iron-containing enzymes.

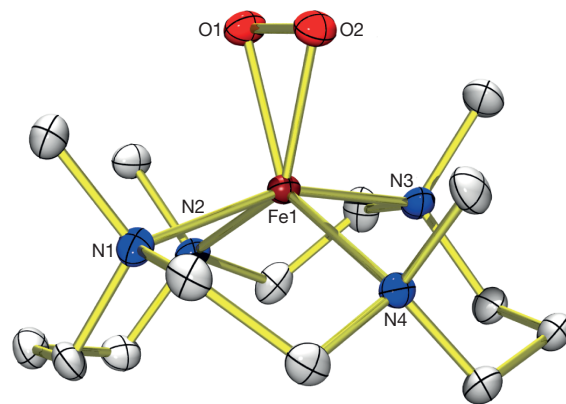
TMC (1,4,8,11-tetramethyl-1,4,8,11-tetraazacyclotetradecane) is a macrocyclic ligand of remarkable versatility in the field of biomimetic chemistry of dioxygen activation by metal complexes. A variety of metal complexes of superoxo, peroxo and oxo ligands showing a wide range of properties have been recently synthesized and characterized using the TMC ligand<sup>8–11</sup>. In the case of iron, TMC complexes of oxo and peroxo ligands are known<sup>11,12</sup>; however, neither the structure of the latter nor that of any other iron(III)–peroxo complexes have been reported before this study. We report here the X-ray crystal structure of a high-spin iron(III)–peroxo complex bearing the TMC ligand,  $[\text{Fe}(\text{III})(\text{TMC})(\text{OO})]^+$  (**1**; Fig. 1).

The iron(III)–peroxo complex, prepared by reacting  $[\text{Fe}(\text{II})(\text{TMC})(\text{CF}_3\text{SO}_3)_2]$  with 5 equiv.  $\text{H}_2\text{O}_2$  in the presence of 2 equiv. triethylamine in  $\text{CF}_3\text{CH}_2\text{OH}$  at 0 °C (ref. 12), was characterized with ultraviolet–visible absorption spectroscopy (Fig. 2a), electrospray ionization mass spectrometry (Supplementary Fig. 1), and electron paramagnetic resonance (EPR) spectroscopy (Supplementary Fig. 2), as reported previously<sup>12</sup>, as well as resonance Raman spectroscopy (Supplementary Fig. 3) and X-ray absorption spectroscopy/extended X-ray absorption fine structure (XAS/EXAFS) (Fig. 2b, c). The X-ray crystal structure of **1**–( $\text{ClO}_4$ ) revealed the mononuclear side-on 1:1 iron complex with  $\text{O}_2$ ; the Fe is in a distorted octahedral geometry, which arises from the triangular  $\text{FeOO}$  moiety with a small O–Fe–O

(‘bite’) angle of  $45.03(17)^\circ$  (Fig. 1; Supplementary Tables 1 and 2). The  $\text{FeOO}$  geometry is similar to the crystallographically characterized structure of naphthalene dioxygenase, where dioxygen binds side-on close to the mononuclear iron at the active site (1.75 Å resolution, O–O distance  $\sim 1.45$  Å)<sup>6</sup>. The structurally determined O–O bond length of 1.463(6) Å in our complex is indicative of peroxo character of the  $\text{OO}$  group<sup>13</sup>, as supported by resonance Raman data<sup>14,15</sup> (Supplementary Figs 3 and 4). It is worth noting that both the O–O bond length and the average Fe–O bond length (1.910 Å) of **1** are longer than those of other metal(III)–peroxo complexes bearing a series of TMC ligands<sup>8–10</sup>. The structure of **1** was further supported by EXAFS analysis (Fig. 2c), which identifies a 2:4 split first shell, with two short Fe–O paths at 1.92 Å and four longer Fe–N TMC paths at  $\sim 2.22$  Å (Supplementary Table 3, fit 1–3; Supplementary Table 4).

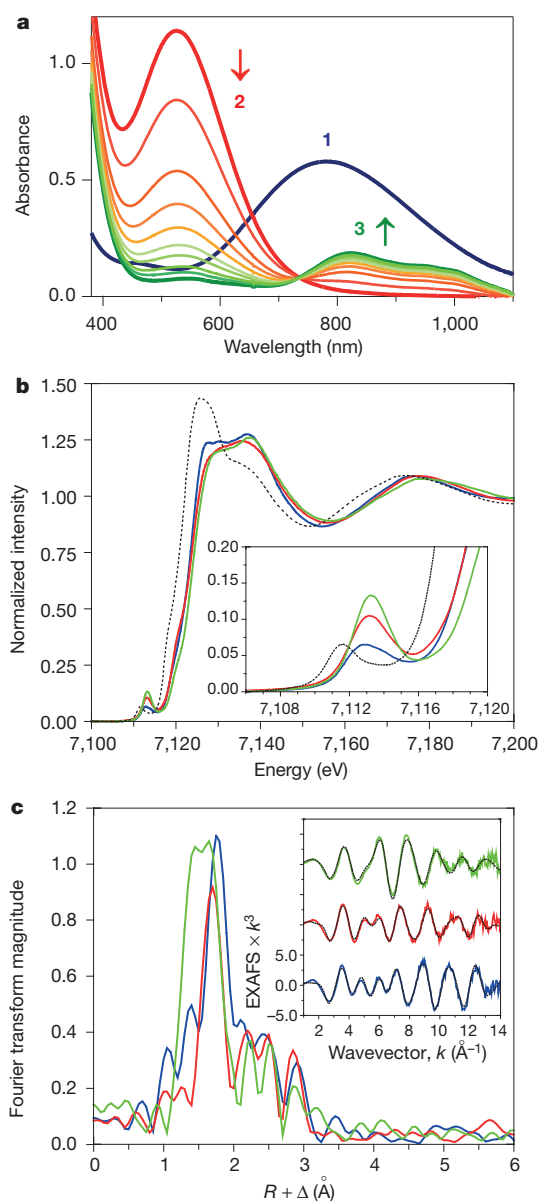
From the structure of **1** (Fig. 1), all four *N*-methyl groups point to the same side of the peroxo moiety, as observed in other metal(III)–peroxo complexes<sup>8–10</sup>. In the case of an iron(IV)–oxo complex bearing the TMC ligand, the *N*-methyl groups of a  $\text{Sc}^{3+}$ -bound  $[\text{Fe}(\text{IV})(\text{TMC})(\text{O})]^{2+}$  complex are also *syn* to the oxo ligand<sup>16</sup>, whereas those in  $[\text{Fe}(\text{IV})(\text{TMC})(\text{O})(\text{CH}_3\text{CN})]^{2+}$  are *anti* to the oxo ligand<sup>11</sup>. In addition, no axial ligand binds to the iron ion *trans* to the peroxo ligand in **1**, which is different from the  $[\text{Fe}(\text{IV})(\text{TMC})(\text{O})(\text{CH}_3\text{CN})]^{2+}$  complex<sup>11</sup> but similar to other metal(III)–peroxo complexes<sup>8–10</sup> as well as the  $\text{Sc}^{3+}$ -bound  $[\text{Fe}(\text{IV})(\text{TMC})(\text{O})]^{2+}$  complex<sup>16</sup>.

Addition of a slight excess of  $\text{HClO}_4$  (for example, 3 equiv. to **1**) to a solution of **1** in acetone/ $\text{CF}_3\text{CH}_2\text{OH}$  (3:1) at  $-40$  °C immediately generated a violet intermediate (**2**) with an electronic absorption band



**Figure 1 | X-ray crystal structure of **1**.** Structure of  $[\text{Fe}(\text{TMC})(\text{OO})]^+$  (**1**), with thermal ellipsoids drawn at the 30% probability level, produced using ORTEP software. Hydrogen atoms are omitted for clarity. Selected bond lengths (Å): Fe–O1 1.906(4), Fe–O2 1.914(4), Fe–N1 2.192(4), Fe–N2 2.256(5), Fe–N3 2.180(5), Fe–N4 2.273(4), O1–O2 1.463(6). Selected angles (°): O1–Fe–O2 45.03(17), Fe–O1–O2 67.8(2), Fe–O2–O1 67.2(2).

<sup>1</sup>Department of Bioinspired Science, Department of Chemistry and Nano Science, Center for Biomimetic Systems, Ewha Womans University, Seoul 120-750, Korea. <sup>2</sup>Department of Chemistry, Stanford University, Stanford, California 94305, USA. <sup>3</sup>Department of Chemistry and Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109, USA. <sup>4</sup>Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford University, Menlo Park, California 94025-7015, USA. <sup>5</sup>Department of Chemistry and Biochemistry, UCLA, Los Angeles, California 90095-1569, USA.



**Figure 2 | Ultraviolet-visible spectra and XAS data of 1, 2 and 3.** In a–c, data for 1, 2 and 3 are shown respectively in blue, red and green. **a**, Ultraviolet-visible spectra of 1, 2 and 3; arrows indicate spectral changes for the conversion of 2 to 3 in the reaction of 1 (1 mM) and 3 equiv. HClO<sub>4</sub> in acetone/CF<sub>3</sub>CH<sub>2</sub>OH (3:1) at -40 °C. **b**, Main panel, Fe K-edge XAS data; inset, expanded pre-edge region. Dotted black line shows starting material, high-spin [Fe(II)(TMC)]<sup>2+</sup>, for reference. **c**, Main panel, Fourier transform of EXAFS data ( $k = 2$ –16).  $R$ , bond length;  $\Delta$ , phase shift of the scattered wave. Inset, EXAFS data (solid lines) with final fits (dashed lines); y axis shows EXAFS intensity multiplied by  $k^3$ . These data show striking differences across the series, most of which are the result of changes to the first coordination sphere.

at  $\lambda_{\text{max}} = 526$  nm, followed by a conversion ( $t_{1/2} \approx 60$  min) from 2 to the corresponding iron(IV)-oxo complex, [Fe(IV)(TMC)(O)]<sup>2+</sup> (3), with an isosbestic point at 735 nm (Fig. 2a; also see Supplementary Fig. 5 for the electrospray ionization mass spectrometry of 3). (Here  $\lambda_{\text{max}}$  is the wavelength of maximum absorption, and  $t_{1/2}$  is the half-life.) Intermediate 2 rapidly reverts back to 1 on addition of 3 equiv. tetramethylammonium hydroxide, suggesting that 1 and 2 are interconverted through the previously reported acid-base chemistry between iron(III)-peroxo and iron(III)-hydroperoxo species<sup>14,17,18</sup>.

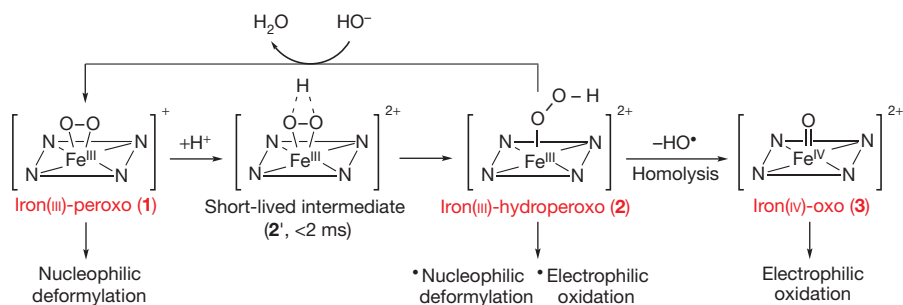
Because the interconversions were fast, the reactions were followed using a stopped-flow spectrometer. On addition of 3 equiv. HClO<sub>4</sub> to 1 in acetone/CF<sub>3</sub>CH<sub>2</sub>OH (3:1) at -40 °C, the absorption band of 1 at

$\lambda_{\text{max}} = 782$  nm disappeared immediately ( $< 2$  ms), but the absorption band of 2 at  $\lambda_{\text{max}} = 526$  nm appeared gradually. The full formation of 2 took  $\sim 100$  ms with an observed rate ( $k_{\text{obs}}$ ) of  $93 \text{ s}^{-1}$  at -40 °C (Supplementary Fig. 6). These spectral changes indicate that another intermediate (2') exists in the conversion of 1 to 2, with an extremely fast conversion of 1 to 2' followed by the relatively slow (that is,  $\sim 100$  ms) conversion of 2' to 2 (Fig. 3 and Supplementary Fig. 7). 2' is likely to be a side-on iron(III)-hydroperoxo species (Fig. 3 and Supplementary Fig. 7), although we have been unable to characterize it spectroscopically owing to its extremely short lifetime. The reverse reaction, which is the conversion of 2 to 1, was also investigated using stopped-flow methods. Addition of 3 equiv. of tetramethylammonium hydroxide to a solution of 2 at -80 °C resulted in the formation of 1 with a clear isosbestic point at 635 nm and a  $k_{\text{obs}}$  of  $19 \text{ s}^{-1}$  (Supplementary Fig. 8; see Fig. 3 for reaction pathways).

Intermediate 2 was characterized using a variety of spectroscopic techniques, including EPR, XAS/EXAFS and resonance Raman. The EPR spectrum of a frozen acetone/CF<sub>3</sub>CH<sub>2</sub>OH (3:1) solution measured at 10 K exhibits signals at  $g = 6.8, 5.2$  and  $1.96$  (Supplementary Fig. 9), which are indicative of a high-spin ( $S = 5/2$ ) Fe(III) species<sup>18,19</sup>. The EXAFS data for 2, compared to 1 (Fig. 2c), also exhibit a distinct shift in the coordination environment, from a 2:4 O:N split first shell in 1 to only a single Fe–O path at a distance of  $1.85 \text{ \AA}$  in 2 (Supplementary Table 3, fit 2-2). This conversion is evident, as a 2:4 split first shell for 2 produces unreasonable bond variances (Supplementary Table 3, fit 2-1). In addition, the remaining Fe–N paths of the TMC ligand have contracted to an average distance of  $2.17 \text{ \AA}$ , reflecting a decrease in ligation. The Fe K-edge of 2 energetically overlays well with that of 1 (Fig. 2b), consistent with the assignment of 2 as a high-spin Fe(III) system. The  $1s \rightarrow 3d$  Fe K pre-edge feature of 2 exhibits an increase in pre-edge intensity from 17.5 to 25.6 units, 1 to 2, respectively (Supplementary Table 5). A pre-edge intensity of 25.6 units is substantially larger than those of other six-coordinate or even five-coordinate complexes<sup>20</sup>, thus favouring a five-coordinate structure for 2 (Supplementary Fig. 10).

On 531-nm excitation at 77 K, the resonance Raman spectrum of <sup>16</sup>O-labelled 2 in d<sub>6</sub>-acetone shows two isotopically sensitive bands at  $658$  and  $868 \text{ cm}^{-1}$  (Supplementary Fig. 11). The peak at  $658 \text{ cm}^{-1}$  shifts to  $633 \text{ cm}^{-1}$  on <sup>18</sup>O-substitution, and is the Fe–O stretch. The peak at  $868 \text{ cm}^{-1}$  shifts to  $820 \text{ cm}^{-1}$  on <sup>18</sup>O-substitution, and is the O–O stretch. The O–O stretch of 2 is higher in wavenumber than those of other high-spin Fe(III)-OOH(R) complexes (for example,  $830 \text{ cm}^{-1}$  for [Fe(H<sub>2</sub>bppa)(OOH)]<sup>2+</sup>)<sup>19</sup> and is much higher than those of low-spin Fe(III)-OOH(R) complexes (for example,  $790 \text{ cm}^{-1}$  for [Fe(N4Py)(OOH)]<sup>2+</sup>)<sup>21</sup>, consistent with the conclusion that 2 is a high-spin Fe(III)-OOH complex. In addition, the Fe–O stretch of 2 is higher than those of six-coordinate high-spin Fe(III)-OOH(R) complexes (for example,  $621 \text{ cm}^{-1}$  for [Fe(H<sub>2</sub>bppa)(OOH)]<sup>2+</sup>)<sup>19</sup>, indicating a stronger Fe–O bond that would be consistent with the absence of a *trans*-axial ligand in 2. This observation is also consistent with the above XAS results suggesting a five-coordinate model for 2. These spectroscopic results are further supported by density functional theory calculations that indicate that a high-spin [Fe(III)(TMC)(OOH)]<sup>2+</sup> complex with its methyl groups oriented *syn* to the OOH<sup>-</sup> ligand does not bind a *trans*-axial ligand (Supplementary Fig. 12).

Knowing that an iron(IV)-oxo complex, [Fe(IV)(TMC)(O)]<sup>2+</sup> (3), is formed as the decay product of 2 (Fig. 2; Supplementary Figs 5 and 13), two possible mechanisms of hydroperoxide O–O bond cleavage of 2 are considered: one is the heterolytic O–O bond cleavage of the hydroperoxide ligand of 2, which would generate an Fe(V)-oxo species, followed by one-electron reduction of the Fe(V)-oxo species that results in the generation of 3 (pathways A and B in Supplementary Fig. 14). The other possibility is the homolytic hydroperoxide O–O bond cleavage of 2, affording 3 and a hydroxyl radical (pathway C in Supplementary Fig. 14). Recently, the former mechanism has been proposed<sup>22</sup>, based on the observation that the formation rate of 3 from



**Figure 3 | Iron–oxygen intermediates.** Generation, structural and spectroscopic characterization, and reactivities of mononuclear non-haem iron–oxygen intermediates detected in the reactivity studies of **1**.

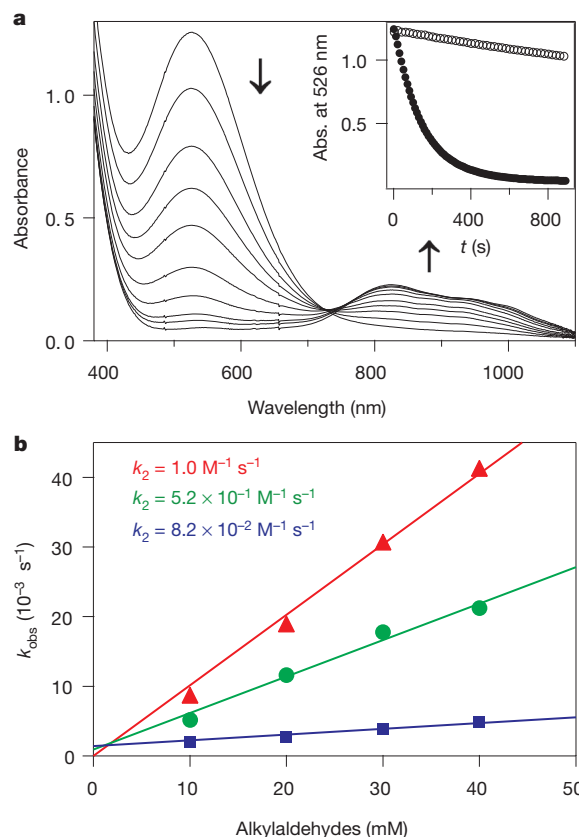
**2** was accelerated by increasing the proton concentration in  $\text{CH}_3\text{CN}$  solution. However, under different solvent conditions, we found no proton concentration effect on the rate of the hydroperoxo O–O bond cleavage of **2**. In fact, the formation rates were essentially the same irrespective of the proton concentration in acetone/ $\text{CF}_3\text{CH}_2\text{OH}$  (3:1) and other solvents except  $\text{CH}_3\text{CN}$  (Supplementary Table 6).

Additional evidence that argues against the formation of an iron(v)–oxo species via O–O bond heterolysis was obtained by carrying out reactions in the presence and absence of substrates (Supplementary Fig. 15). If **3** were in fact the product of the one-electron reduction of an iron(v)–oxo species, then the amount of **3** formed in the presence of the substrates should decrease due to fast reactions between the highly reactive iron(v)–oxo species and the substrates (pathways A and D in Supplementary Fig. 14). However, the amounts of **3** formed in the presence and absence of substrates were the same, implying that an iron(v)–oxo species was not generated via an O–O bond heterolysis mechanism in the course of the formation of **3** from **2**. It is also significant that this chemistry, in which a high-spin iron(III)–hydroperoxo species undergoes O–O bond homolysis, has not been reported previously for other high-spin iron(III)–hydroperoxo species<sup>21,23</sup>. A possible reason is that the  $\text{Fe}(\text{IV})=\text{O}$  product in our reaction is an intermediate-spin, six-coordinate complex. Thus, changes along the reaction coordinate in spin state and coordination number could contribute to the energetics of this O–O bond cleavage.

The detailed reactivities of the three intermediates, iron(III)–peroxo (**1**), iron(III)–hydroperoxo (**2**), and iron(IV)–oxo (**3**), have been investigated in both nucleophilic and electrophilic reactions (Fig. 3). The nucleophilic characters of all three intermediates were tested in aldehyde deformylation reactions<sup>24</sup>. On addition of 2-phenylpropionaldehyde (2-PPA) to **1** and **3** in acetone/ $\text{CF}_3\text{CH}_2\text{OH}$  (3:1) at  $-40^\circ\text{C}$ , the intermediates remained intact without any spectral change (Supplementary Fig. 16). These results indicate that **1** and **3** are relatively unreactive in nucleophilic oxidative reactions at  $-40^\circ\text{C}$ , although **1** showed a reactivity with the aldehyde at high temperature (for example,  $15^\circ\text{C}$ )<sup>12</sup>. In contrast, intermediate **2** reacted rapidly with 2-PPA at  $-40^\circ\text{C}$ , resulting in the disappearance of its characteristic ultraviolet–visible band. This follows a first-order decay profile and forms **3** with an isosbestic point at 735 nm (Fig. 4a; see Supplementary Fig. 17 for aldehyde concentration effect). The high reactivity of **2** in nucleophilic reactions, compared to the side-on iron(III)–peroxo analogue **1**, is ascribed to the end-on binding mode of the hydroperoxo ligand<sup>24,25</sup> and supported by density functional theory calculations (Supplementary Fig. 18). The reactivity of **2** was further investigated using primary ( $1^\circ\text{-CHO}$ ), secondary ( $2^\circ\text{-CHO}$ ), and tertiary ( $3^\circ\text{-CHO}$ ) aldehydes (Fig. 4b), and the observed reactivity order of  $1^\circ\text{-CHO} > 2^\circ\text{-CHO} > 3^\circ\text{-CHO}$  supports the nucleophilic character of **2** (Supplementary Fig. 19 shows additional evidence for the nucleophilic character of **2**).

The electrophilic characters of **1**, **2** and **3** were also investigated in the oxidation of alkylaromatic compounds with weak C–H bonds,

such as xanthene ( $75.5 \text{ kcal mol}^{-1}$ ) and 9,10-dihydroanthracene (DHA,  $77 \text{ kcal mol}^{-1}$ ). **1** did not show any significant spectral change on addition of substrates in acetone/ $\text{CF}_3\text{CH}_2\text{OH}$  (3:1) at  $-20^\circ\text{C}$  (Supplementary Fig. 20a). In contrast, **2** and **3** reacted with DHA under the same conditions (Supplementary Figs 20b and 21a), showing that both **2** and **3** are capable of abstracting a hydrogen atom from DHA and that **2** has a similar reactivity to **3** in this C–H bond activation reaction. Second-order rate constants of  $8.1 \times 10^{-1}$  and  $2.4 \times 10^{-2} \text{ M}^{-1} \text{ s}^{-1}$  were determined in the oxidation of xanthene and DHA, respectively, by **2** at  $-20^\circ\text{C}$  (Supplementary Fig. 21b). On the basis of the above observations—that the reaction rates are dependent



**Figure 4 | Reactivity studies of **2** with aldehydes.** **a**, Ultraviolet–visible spectral changes showing the decay of **2** (1 mM) and formation of **3** on addition of 2-PPA (50 mM). Inset, time courses of the absorbance change of **2** at 526 nm for the reaction with 2-PPA (filled circles) and for the natural decay (open circles). **b**, Second-order rate constants determined in the reactions of **2** with pentanal ( $1^\circ\text{-CHO}$ ; triangles), 2-methylbutanal ( $2^\circ\text{-CHO}$ ; circles), and pivalaldehyde ( $3^\circ\text{-CHO}$ ; squares). Standard deviation is  $<10\%$  of the data used in the plot.

on the substrate concentration and the bond dissociation energies of the substrates (Supplementary Fig. 21b)<sup>26,27</sup>—we conclude that **2** is the active oxidant in abstracting an H atom from the substrates.

In conclusion, iron(III)–hydroperoxo intermediates have been proposed as active oxidants in cytochrome P450-catalysed deformylation of aldehydes<sup>28,29</sup> and activated bleomycin-mediated DNA cleavage via H-atom abstraction<sup>21,30</sup>. We have provided here direct experimental evidence that a high-spin iron(III)–hydroperoxo species is capable of acting as an active oxidant in both nucleophilic and electrophilic reactions.

## METHODS SUMMARY

The iron(III)–peroxo complex, [Fe(III)(TMC)(OO)]<sup>+</sup> (**1**), was prepared by reacting [Fe(II)(TMC)(CF<sub>3</sub>SO<sub>3</sub>)<sub>2</sub>] with 5 equiv. H<sub>2</sub>O<sub>2</sub> in the presence of 2 equiv. triethylamine in CF<sub>3</sub>CH<sub>2</sub>OH at 0 °C. Unlike the highly unstable nature of **1** in CH<sub>3</sub>CN (refs 12, 22), **1** prepared in CF<sub>3</sub>CH<sub>2</sub>OH persisted for several hours at 0 °C, and this greater thermal stability of **1** in alcoholic solvents allowed for the isolation of crystals with ~80% yield, which were used for spectroscopic characterization and reactivity studies. Crystals suitable for structural analysis were obtained from CH<sub>3</sub>OH/diethyl ether with excess NaClO<sub>4</sub> at –40 °C. See experimental section in Supplementary Information for detailed experimental conditions and procedures, spectroscopic and kinetics analyses, and computational calculations.

Received 18 March; accepted 5 September 2011.

- Solomon, E. I. *et al.* Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **100**, 235–350 (2000).
- Kovaleva, E. G. & Lipscomb, J. D. Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nature Chem. Biol.* **4**, 186–193 (2008).
- Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T. & Drennan, C. L. Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440**, 368–371 (2006).
- Rittle, J. & Green, M. T. Cytochrome P450 compound I: capture, characterization, and C–H bond activation kinetics. *Science* **330**, 933–937 (2010).
- Kovaleva, E. G. & Lipscomb, J. D. Crystal structures of Fe<sup>2+</sup> dioxygenase superoxo, alkylperoxo, and bound product intermediates. *Science* **316**, 453–457 (2007).
- Karlsson, A. *et al.* Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. *Science* **299**, 1039–1042 (2003).
- Cicchillo, R. M. *et al.* An unusual carbon–carbon bond cleavage reaction during phosphinothricin biosynthesis. *Nature* **459**, 871–874 (2009).
- Seo, M. S. *et al.* [Mn(tmc)(O<sub>2</sub>)<sup>+</sup>]: a side-on peroxido manganese(III) complex bearing a non-heme ligand. *Angew. Chem. Int. Edn* **46**, 377–380 (2007).
- Cho, J. *et al.* Synthesis, structural, and spectroscopic characterization and reactivities of mononuclear cobalt(III)–peroxo complexes. *J. Am. Chem. Soc.* **132**, 16977–16986 (2010).
- Cho, J. *et al.* Geometric and electronic structure and reactivity of a mononuclear ‘side-on’ nickel(III)–peroxo complex. *Nature Chem.* **1**, 568–572 (2009).
- Rohde, J.-U. *et al.* Crystallographic and spectroscopic characterization of a nonheme Fe(IV)=O complex. *Science* **299**, 1037–1039 (2003).
- Annaraj, J., Suh, Y., Seo, M. S., Kim, S. O. & Nam, W. Mononuclear nonheme ferric-peroxo complex in aldehyde deformylation. *Chem. Commun.* 4529–4531 (2005).
- Cramer, C. J., Tolman, W. B., Theopold, K. H. & Rheingold, A. L. Variable character of O–O and M–O bonding in side-on (η<sup>2</sup>) 1:1 metal complexes of O<sub>2</sub>. *Proc. Natl Acad. Sci. USA* **100**, 3635–3640 (2003).
- Neese, F. & Solomon, E. I. Detailed spectroscopic and theoretical studies on [Fe(EDTA)(O<sub>2</sub>)<sup>3-</sup>]: electronic structure of the side-on ferric-peroxide bond and its relevance to reactivity. *J. Am. Chem. Soc.* **120**, 12829–12848 (1998).
- Liu, J.-G. *et al.* Spectroscopic characterization of a hydroperoxo-heme intermediate: conversion of a side-on peroxo to an end-on hydroperoxo complex. *Angew. Chem. Int. Edn* **48**, 9262–9267 (2009).
- Fukuzumi, S. *et al.* Crystal structure of a metal ion-bound oxoiron(IV) complex and implications for biological electron transfer. *Nature Chem.* **2**, 756–759 (2010).
- Jensen, K. B., McKenzie, C. J., Nielsen, L. P., Pedersen, J. Z. & Svendsen, H. M. Deprotonation of low-spin mononuclear iron(III)–hydroperoxide complexes give transient blue species assigned to high-spin iron(III)–peroxide complexes. *Chem. Commun.* 1313–1314 (1999).
- Simaan, A. J., Banse, F., Girerd, J.-J., Wieghardt, K. & Bill, E. The electronic structure of non-heme iron(III)–hydroperoxo and iron(III)–peroxo model complexes studied by Mössbauer and electron paramagnetic resonance spectroscopies. *Inorg. Chem.* **40**, 6538–6540 (2001).
- Wada, A. *et al.* Reactivity of hydroperoxide bound to a mononuclear non-heme iron site. *Inorg. Chem.* **41**, 616–618 (2002).
- Westre, T. E. *et al.* A multiplet analysis of Fe K-edge 1s → 3d pre-edge features of iron complexes. *J. Am. Chem. Soc.* **119**, 6297–6314 (1997).
- Lehnert, N., Neese, F., Ho, R. Y. N., Que, L. Jr & Solomon, E. I. Electronic structure and reactivity of low-spin Fe(III)–hydroperoxo complexes: comparison to activated bleomycin. *J. Am. Chem. Soc.* **124**, 10810–10822 (2002).
- Li, F. *et al.* Characterization of a high-spin non-heme Fe<sup>III</sup>–OOH intermediate and its quantitative conversion to an Fe<sup>IV</sup>=O complex. *J. Am. Chem. Soc.* **133**, 7256–7259 (2011).
- Shan, X. *et al.* X-ray absorption spectroscopic studies of high-spin nonheme (alkylperoxo)iron(III) intermediates. *Inorg. Chem.* **46**, 8410–8417 (2007).
- Wertz, D. L. & Valentine, J. S. Nucleophilicity of iron–peroxo porphyrin complexes. *Struct. Bonding* **97**, 37–60 (2000).
- Selke, M. & Valentine, J. S. Switching on the nucleophilic reactivity of a ferric porphyrin peroxo complex. *J. Am. Chem. Soc.* **120**, 2652–2653 (1998).
- Mayer, J. M. Understanding hydrogen atom transfer: from bond strengths to Marcus theory. *Acc. Chem. Res.* **44**, 36–46 (2011).
- Sastri, C. V. *et al.* Axial ligand tuning of a nonheme iron(IV)–oxo unit for hydrogen atom abstraction. *Proc. Natl Acad. Sci. USA* **104**, 19181–19186 (2007).
- Vaz, A. D. N., Roberts, E. S. & Coon, M. J. Olefin formation in the oxidative deformylation of aldehydes by cytochrome P-450. Mechanistic implications for catalysis by oxygen-derived peroxide. *J. Am. Chem. Soc.* **113**, 5886–5887 (1991).
- Akhtar, M., Corina, D., Miller, S., Shyadehi, A. Z. & Wright, J. N. Mechanism of the acyl-carbon cleavage and related reactions catalyzed by multifunctional P-450s: studies on cytochrome P-450<sub>17α</sub>. *Biochemistry* **33**, 4410–4418 (1994).
- Solomon, E. I., Wong, S. D., Liu, L. V., Decker, A. & Chow, M. S. Peroxo and oxo intermediates in mononuclear nonheme iron enzymes and related active sites. *Curr. Opin. Chem. Biol.* **13**, 99–113 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The work was supported by NRF/MEST of Korea through the CRI (W.N.), the GRL (2010-00353; W.N.), the WCU (R31-2008-000-10010-0; W.N. and J.S.V.) and the 2011 KRIST OASIS Project (W.N.), by NIH grants GM 40392 (E.I.S.) and RR-001209 (K.O.H.), and by NSF grant MCB 0919027 (E.I.S.). J.J.B. acknowledges a Fellowship from NSF EAPSI (OISE-1014685) and the Warner Linfield Award from the University of Michigan. SSRL operations are funded by the Department of Energy (DOE) Office of Science and operated by Stanford University. The SSRL Structural Molecular Biology programme is supported by the DOE, Office of Biological and Environmental Research, and by the NIH, National Center for Research Resources (grant 5P41RR001209), Biomedical Technology Program.

**Author Contributions** J.C., J.S.V., E.I.S. and W.N. conceived and designed the experiments. J.C., S.J., S.A.W., L.V.L., E.A.K., J.J.B., M.H.L., B.H. and K.O.H. performed the experiments and analysed the data. J.C., S.A.W., L.V.L., J.S.V., E.I.S. and W.N. co-wrote the Letter.

**Author Information** The crystallographic data for **1** have been deposited with the Cambridge Crystallographic Data Center under accession number CCDC 804038. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.S.V. ([jsv@chem.ucla.edu](mailto:jsv@chem.ucla.edu)), E.I.S. ([edward.solomon@stanford.edu](mailto:edward.solomon@stanford.edu)) or W.N. ([wnam@ewha.ac.kr](mailto:wnam@ewha.ac.kr)).

# A draft genome of *Yersinia pestis* from victims of the Black Death

Kirsten I. Bos<sup>1\*</sup>, Verena J. Schuenemann<sup>2\*</sup>, G. Brian Golding<sup>3</sup>, Hernán A. Burbano<sup>4</sup>, Nicholas Waglechner<sup>5</sup>, Brian K. Coombes<sup>5</sup>, Joseph B. McPhee<sup>5</sup>, Sharon N. DeWitte<sup>6,7</sup>, Matthias Meyer<sup>4</sup>, Sarah Schmedes<sup>8</sup>, James Wood<sup>9</sup>, David J. D. Earn<sup>5,10</sup>, D. Ann Herring<sup>11</sup>, Peter Bauer<sup>12</sup>, Hendrik N. Poinar<sup>1,3,5</sup> & Johannes Krause<sup>2,12</sup>

Technological advances in DNA recovery and sequencing have drastically expanded the scope of genetic analyses of ancient specimens to the extent that full genomic investigations are now feasible and are quickly becoming standard<sup>1</sup>. This trend has important implications for infectious disease research because genomic data from ancient microbes may help to elucidate mechanisms of pathogen evolution and adaptation for emerging and re-emerging infections. Here we report a reconstructed ancient genome of *Yersinia pestis* at 30-fold average coverage from Black Death victims securely dated to episodes of pestilence-associated mortality in London, England, 1348–1350. Genetic architecture and phylogenetic analysis indicate that the ancient organism is ancestral to most extant strains and sits very close to the ancestral node of all *Y. pestis* commonly associated with human infection. Temporal estimates suggest that the Black Death of 1347–1351 was the main historical event responsible for the introduction and widespread dissemination of the ancestor to all currently circulating *Y. pestis* strains pathogenic to humans, and further indicates that contemporary *Y. pestis* epidemics have their origins in the medieval era. Comparisons against modern genomes reveal no unique derived positions in the medieval organism, indicating that the perceived increased virulence of the disease during the Black Death may not have been due to bacterial phenotype. These findings support the notion that factors other than microbial genetics, such as environment, vector dynamics and host susceptibility, should be at the forefront of epidemiological discussions regarding emerging *Y. pestis* infections.

The Black Death of 1347–1351, caused by the bacterium *Yersinia pestis*<sup>2,3</sup>, provides one of the best historical examples of an emerging infection with rapid dissemination and high mortality, claiming an estimated 30–50% of the European population in only a five-year period<sup>4</sup>. Discrepancies in epidemiological trends between the medieval disease and modern *Y. pestis* infections have ignited controversy over the pandemic's aetiological agent<sup>5,6</sup>. Although ancient DNA investigations have strongly implicated *Y. pestis*<sup>2,3</sup> in the ancient pandemic, genetic changes in the bacterium may be partially responsible for differences in disease manifestation and severity. To understand the organism's evolution it is necessary to characterize the genetic changes involved in its transformation from a sylvatic pathogen to one capable of pandemic human infection on the scale of the Black Death, and to determine its relationship with currently circulating strains. Here we begin this discussion by presenting the first draft genome sequence of the ancient pathogen.

*Y. pestis* is a recently evolved descendent of the soil-dwelling bacillus *Yersinia pseudotuberculosis*<sup>7</sup>, which in the course of its evolution

acquired two additional plasmids (pMT1 and pPCP1) that provide it with specialized mechanisms for infiltrating mammalian hosts. To investigate potential evolutionary changes in one of these plasmids, we reported on the screening of 46 teeth and 53 bones from the East Smithfield collection of London, England for presence of the *Y. pestis*-specific pPCP1 (ref. 3). Historical data indicate that the East Smithfield burial ground was established in late 1348 or early 1349 specifically for interment of Black Death victims<sup>8</sup> (Supplementary Figs 1 and 2), making the collection well-suited for genetic investigations of ancient *Y. pestis*. DNA sequence data for five teeth obtained via molecular capture of the full *Y. pestis*-specific pPCP1 revealed a C to T damage pattern characteristic of authentic endogenous ancient DNA<sup>9</sup>, and assembly of the pooled Illumina reads permitted the reconstruction of 98.68% of the 9.6-kilobase plasmid at a minimum of twofold coverage<sup>3</sup>.

To evaluate the suitability of capture-based methods for reconstructing the complete ancient genome, multiple DNA extracts from both roots and crowns stemming from four of the five teeth which yielded the highest pPCP1 coverage<sup>3</sup> were used for array-based enrichment (Agilent) and subsequent high-throughput sequencing on the Illumina GAII platform<sup>10</sup>. Removal of duplicate molecules and subsequent filtering produced a total of 2,366,647 high quality chromosomal reads (Supplementary Table 1a, b) with an average fragment length of 55.53 base pairs (Supplementary Fig. 4), which is typical for ancient DNA. Coverage estimates yielded an average of 28.2 reads per site for the chromosome, and 35.2 and 31.2 for the pCD1 and pMT1 plasmids, respectively (Fig. 1a, c, d and Supplementary Table 1b, c). Coverage was predictably low for pPCP1 (Fig. 1e) because probes specific to this plasmid were not included on the arrays. Coverage correlated with GC content (Supplementary Fig. 6), a trend previously observed for high-throughput sequence data<sup>11</sup>. The coverage on each half of the chromosome was uneven due to differences in sequencing depth between the two arrays, with 36.46 and 22.41 average reads per site for array 1 and array 2, respectively. Although greater depth contributed to more average reads per site, it did not increase overall coverage, with both arrays covering 93.48% of the targeted regions at a minimum of onefold coverage (Supplementary Table 1b). This indicates that our capture procedure successfully retrieved template molecules from all genomic regions accessible via this method, and that deeper sequencing would not result in additional data for CO92 template regions not covered in our data set.

Genome architecture is known to vary widely among extant *Y. pestis* strains<sup>12</sup>. To extrapolate gene order in our ancient genome, we analysed reads mapping to the CO92 reference for all extracts stemming

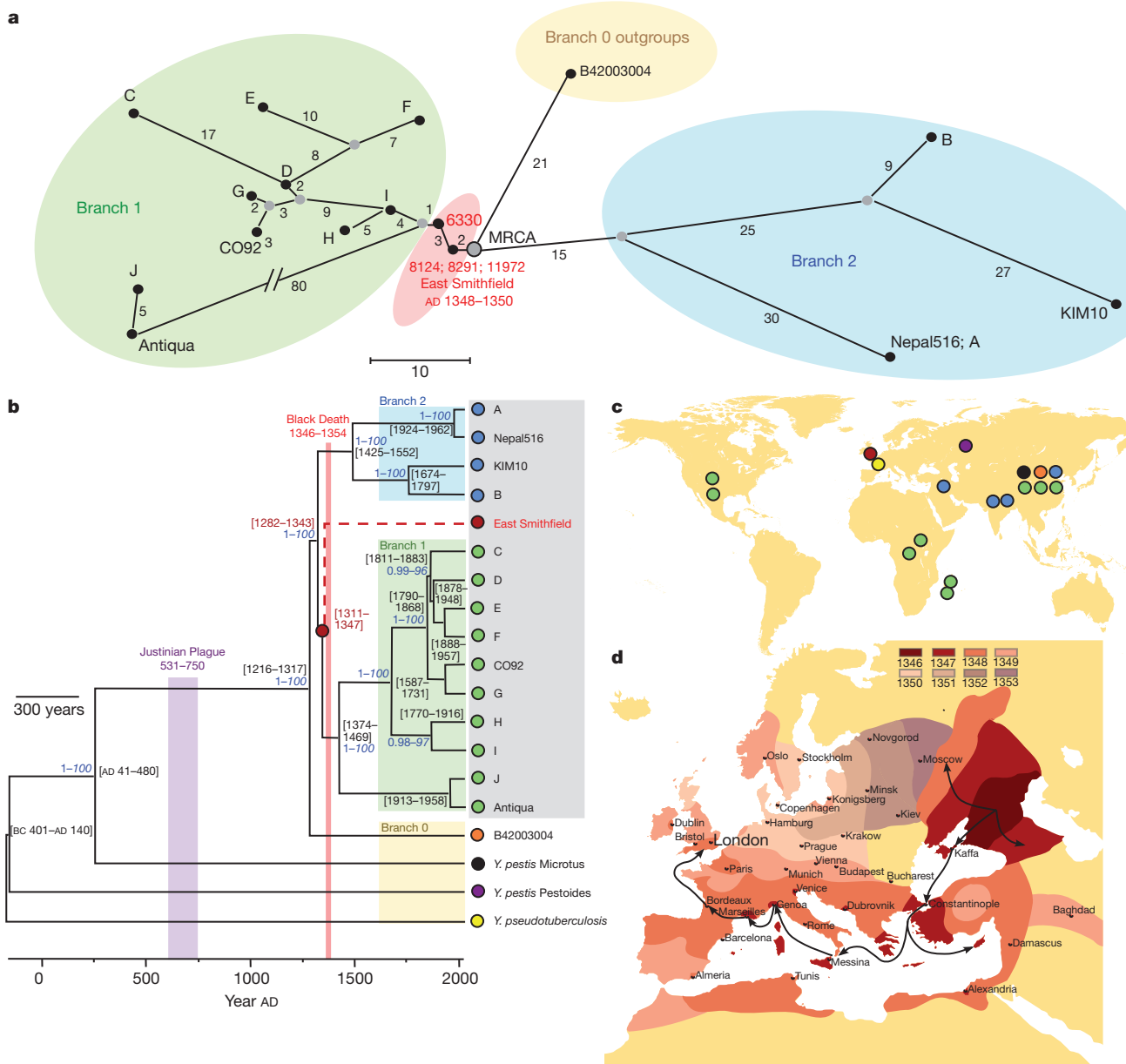
<sup>1</sup>McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L8, Canada. <sup>2</sup>Institute for Archaeological Sciences, Rümelinstr. 23, University of Tübingen, 72070 Tübingen, Germany. <sup>3</sup>Biology Department, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L8, Canada. <sup>4</sup>Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>5</sup>Michael G. DeGroot Institute for Infectious Disease Research, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L8, Canada. <sup>6</sup>Department of Anthropology, University of South Carolina, Columbia, South Carolina 29208, USA. <sup>7</sup>Department of Biological Sciences, University of South Carolina, Columbia, South Carolina 29208, USA. <sup>8</sup>Institute of Applied Genetics, University of North Texas Health Science Center, 3500 Camp Bowie Boulevard, Fort Worth, Texas 76107, USA. <sup>9</sup>Department of Anthropology and Population Research Institute, Pennsylvania State University, University Park, Pennsylvania 16802, USA. <sup>10</sup>Department of Mathematics and Statistics, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada. <sup>11</sup>Department of Anthropology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L8, Canada. <sup>12</sup>Human Genetics Department, Medical Faculty, University of Tübingen, 72070 Tübingen, Germany.

\*These authors contributed equally to this work.

derived from deaminated cytosines that would have been removed in the current investigation via uracil-DNA-glycosylase treatment before array capture.

To place our ancient genome in a phylogenetic context, we characterized all 1,694 previously identified phylogenetically informative positions<sup>14</sup> (Supplementary Table 4), and compared those from our ancient organism against aggregate base call data for 17 publicly available *Y. pestis* genomes and the ancestral *Y. pseudotuberculosis*. When considered separately, sequences from three of the four victims fall only two substitutions from the root of all extant human pathogenic *Y. pestis* strains (Fig. 3a), and they show a closer relationship to branch 1 *Y. pestis* than to branch 2; however, one of the four victims (individual 6330) was infected with a strain that contained three additional derived positions seen in all other branch 1 genomes<sup>14</sup>. This suggests either the presence of

multiple strains in the London 1348–1350 pandemic or microevolutionary changes accruing in one strain, which is known to occur in disease outbreaks<sup>15</sup>. Additional support for *Y. pestis* microevolution is indicated by the presence of several variant positions for which sequence data from one individual shows two different nucleotides at comparable frequencies (Supplementary Table 5). Position 2896636, for example, is a known polymorphic position in extant *Y. pestis* populations<sup>14</sup>, and this position shows the fixed derived state in one individual (6330) and the polymorphic state in another (individual 8291) at minimum fivefold coverage (Supplementary Fig. 7). This provides a remarkable example of microevolution captured during an historical pandemic. The remaining variance positions are unchanged in the 18 extant *Yersinia* genomes, thus they may be unique to the ancient organism and are, therefore, of further interest. Additional sampling



**Figure 3 | Phylogenetic placement and historical context for the East Smithfield strain.** **a**, Median network of ancient and modern *Y. pestis* based on 1,694 variant positions in modern genomes<sup>14</sup>. Coloured circles represent different clades as defined in ref. 13. Gray circles represent hypothetical nodes. **b**, Phylogenetic tree using 1,694 variable positions. Divergence time intervals are shown in calendar years, with neighbour-joining bootstrap support (blue italic) and Bayesian posterior probability (blue). Grey box indicates known

human pathogenic strains. A, NZ ACNQ01000; Nepal516, NC 008149; KIM10, NC 004088; B, NZ AAYT01000; C, NZ ABAT01000; D, NZ ACNS01000; E, NZ AAYS01000; F, NZ AAOS02000; CO92, NC 003143; G, NZ ABCD01000; H, NZ AAYV01000; I, NC 014029; J, NZ AAYR01000; Antiqua, NC 008150. **c**, Geographical origin of genome sequences used in **a** and **b**. **d**, Geographical spread of the Black Death from infection routes reported in ref. 4.

of ancient genomes will assist in determining the frequency of these mutations in co-circulating *Y. pestis* strains, and will clarify the emergence of branch 2 strains that are as yet unreported in ancient samples.

Consistent tree topologies were produced using several construction methods and all major nodes were supported by posterior probability (pp) values of >0.96 and bootstrap values >90 (Fig. 3b and Supplementary Figs 8 and 9). The trees place the East Smithfield sequence close to the ancestral node of all extant human pathogenic *Y. pestis* strains (only two differences in 1,694 positions) and at the base of branch 1 (Fig. 3b). A secure date for the East Smithfield site of 1348–1350 allowed us to assign a tip calibration to the ancient sequence and thus date the divergence time of the modern genomes and the East Smithfield genome using a Bayesian approach. Temporal estimates indicate that all *Y. pestis* commonly associated with human infection shared a common ancestor sometime between 668 and 729 years ago (AD 1282–1343, 95% highest probability density, HPD), encompassing a much smaller time interval than recently published estimates<sup>14</sup> and further indicating that all currently circulating branch 1 and branch 2 isolates emerged during the thirteenth century at the earliest (Fig. 3b), potentially stemming from an Eastern Asian source as has been previously suggested<sup>14</sup>. This implies that the medieval plague was the main historical event that introduced human populations to the ancestor of all known pathogenic strains of *Y. pestis*. This further questions the aetiology of the sixth to eighth century Plague of Justinian, popularly assumed to have resulted from the same pathogen: our temporal estimates imply that the pandemic was either caused by a *Y. pestis* variant that is distinct from all currently circulating strains commonly associated with human infections, or it was another disease altogether.

Although our approach of using an extant *Y. pestis* reference template for bait design precluded our ability to identify genomic regions that may have been present in the ancient organism and were subsequently lost in CO92, genomic comparisons of our ancient sequence against its closest outgroups may yield valuable insights into *Y. pestis* evolution. The *Microtus* 91001 strain is the closest branch 1 and branch 2 relative confirmed to be non-pathogenic to humans<sup>16</sup>, hence genetic changes may represent contributions to the pathogen's adaptation to a human host. Comparisons against this outgroup revealed 113 changes (Supplementary Table 6a, b), many of which are found in genes affecting virulence-associated functions like biofilm formation (*hmsT*), iron-acquisition (*iucD*) or adaptation to the intracellular environment (*phoP*). Similarly, although its virulence potential in humans has yet to be confirmed to our knowledge, *Y. pestis* B42003004 isolated from a Chinese marmot population<sup>17</sup> has been identified as the strain closest to the ancestral node of all *Y. pestis* commonly associated with human plague, and thus may provide key information regarding the organism's evolution. Full genome comparison against the East Smithfield sequence revealed only eight single-nucleotide differences (Supplementary Table 6c), six of which result in non-synonymous changes (Supplementary Table 6d). Although these differences probably do not affect virulence, the influence of gene loss, gene gain or genetic rearrangements, all of which are well documented in *Y. pestis*<sup>12,18</sup>, is as yet undetermined. In more recent evolutionary terms, single-nucleotide differences in several known pathogenicity-associated genes were found between our ancient genome and the CO92 reference sequence (Supplementary Table 3), which may represent further adaptations to human hosts.

Through enrichment by DNA capture coupled with targeted high throughput DNA sequencing, we have reconstructed a draft genome for what is arguably the most devastating human pathogen in history, and revealed that the medieval plague of the fourteenth century was probably responsible for its introduction and widespread distribution in human populations. This indicates that the pathogen implicated in the Black Death has close relatives in the twenty-first century that are both endemic and emerging<sup>19</sup>. Introductions of new pathogens to populations are often associated with increased incidence and severity

of disease<sup>20</sup> and although the mechanisms governing this phenomenon are complex<sup>21</sup>, genetic data from ancient infectious diseases will provide invaluable contributions towards our understanding of host–pathogen coevolution. The Black Death is a seminal example of an emerging infection, travelling across Europe and claiming the lives of an estimated 30 million people in only 5 years, which is much faster than contemporary rates of bubonic or pneumonic plague infection<sup>22</sup> and dissemination<sup>7,8</sup>. Regardless, although no extant *Y. pestis* strain possesses the same genetic profile as our ancient organism, our data suggest that few changes in known virulence-associated genes have accrued in the organism's 660 years of evolution as a human pathogen, further suggesting that its perceived increased virulence in history<sup>23</sup> may not be due to novel fixed point mutations detectable via the analytical approach described here. At our current resolution, we posit that molecular changes in pathogens are but one component of a constellation of factors contributing to changing infectious disease prevalence and severity, where genetics of the host population<sup>24</sup>, climate<sup>25</sup>, vector dynamics<sup>26</sup>, social conditions<sup>27</sup> and synergistic interactions with concurrent diseases<sup>28</sup> should be foremost in discussions of population susceptibility to infectious disease and host–pathogen relationships with reference to *Y. pestis* infections.

## METHODS SUMMARY

DNA from dental pulp was extracted and converted into sequencing libraries as previously described<sup>3</sup>. Potential sequencing artefacts resulting from deaminated nucleotides were eliminated by treatment of the DNA extracts with uracil-DNA-glycosylase and endonuclease VIII. DNA extracts were subsequently converted into sequencing libraries and amplified to incorporate unique sequence tags on both ends of the molecule. Two Agilent DNA capture arrays were designed for capture of the full *Y. pestis* chromosome (4.6 megabases), and the pCD1 (70 kb) and pMT1 (100 kb) plasmids using the modern *Y. pestis* strain CO92 (accession numbers NC\_003143, NC\_003131, NC\_003134) for bait design with 3 bp tiling density. Serial array capture was performed over two copies of each array using the enriched fraction from the first round of capture as a template for a second round. The resulting products were amplified and pooled in equimolar amounts. All templates were sequenced for 76 cycles from both ends on the Illumina GAII platform, and reads merged into single fragments were included in subsequent analyses only if forward and reverse sequences overlapped by a minimum of 11 bp. Reads were mapped against the CO92 genome using the software BWA, and molecules with the same start and end coordinates were removed with the *rmDup* program in the *samtools* suite. Reference-guided sequence assembly was performed using Velvet version 1.1.03, with mapped and unmapped reads supplied in separate channels. Single-nucleotide differences were determined at a minimum of fivefold coverage and base frequency of at least 95% for both a pooled data set for all individuals and one in which all individuals were treated separately. A median network was constructed on these base calls using SplitsTree4. Phylogenetic trees were constructed using parsimony, neighbour-joining (MEGA 4.1) and Bayesian methods, and coalescence dates were determined in BEAST using both a strict and a relaxed molecular clock (Supplementary Fig. 9).

Received 25 July; accepted 9 September 2011.

Published online 12 October 2011.

1. Stoneking, M. & Krause, J. Learning about human population history from ancient and modern genomes. *Nature Rev. Genet.* **12**, 603–614 (2011).
2. Haensch, S. *et al.* Distinct clones of *Yersinia pestis* caused the Black Death. *PLoS Pathog.* **6**, e1001134 (2010).
3. Schuenemann, V. J. *et al.* Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of *Yersinia pestis* from victims of the Black Death. *Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.1105107108 (29 August, 2011).
4. Benedictow, O. J. *The Black Death 1346–1353: The Complete History* (Boydell Press, 2004).
5. Scott, S. & Duncan, C. J. *Biology of Plagues* (Cambridge Univ. Press, 2001).
6. Cohn, S. K. *The Black Death Transformed: Disease and Culture in Early Renaissance Europe* (Arnold Publication, 2003).
7. Achtman, M. *et al.* *Yersinia pestis*, the cause of the plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* **96**, 14043–14048 (1999).
8. Cowal, L., Grainger, I., Hawkins, D. & Mikulski, R. 2008. The Black Death Cemetery, East Smithfield, London (Museum of London Archaeology Service, 2008).
9. Briggs, A. W. *et al.* Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl Acad. Sci. USA* **104**, 14616–14621 (2007).

10. Hodges, E. *et al.* Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. *Nature Protocols* **4**, 960–974 (2009).
11. Green, R. E. *et al.* A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134**, 416–426 (2008).
12. Chain, P. S. G. *et al.* Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* **101**, 13826–13831 (2004).
13. Achtman, M. *et al.* Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc. Natl Acad. Sci. USA* **101**, 17837–17842 (2004).
14. Morelli, G. *et al.* *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nature Genet.* **42**, 1140–1143 (2010).
15. Harris, S. R. *et al.* Evolution of MRSA during hospital transmission and intercontinental spread. *Science* **327**, 469–474 (2010).
16. Song, Y. *et al.* Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to humans. *DNA Res.* **11**, 179–197 (2004).
17. Eppinger, M. *et al.* Draft genome sequences of *Yersinia pestis* isolates from natural foci of endemic plague in China. *J. Bacteriol.* **191**, 7628–7629 (2009).
18. Pouillot, F., Fayolle, C. & Carniel, E. Characterization of chromosomal regions conserved in *Yersinia pseudotuberculosis* and lost by *Yersinia pestis*. *Infect. Immun.* **76**, 4592–4599 (2008).
19. Stenseth, N. C. *et al.* Plague: past, present, and future. *PLoS Med.* **5**, e3 (2008).
20. Baum, J. & Bar-Gal, G. K. in *Emerging Pathogens Archaeology, Ecology & Evolution of Infectious Diseases* (eds Greenblat, C. L. & Spigleman, M.) 67–78 (Oxford Univ. Press, 2003).
21. Brown, N. F. *et al.* Crossing the line: selection and evolution of virulence traits. *PLoS Pathog.* **2**, e42 (2006).
22. WHO. Interregional meeting on prevention and control of plague. ([http://www.who.int/csr/resources/publications/WHO\\_HSE\\_EPR\\_2008\\_3w.pdf](http://www.who.int/csr/resources/publications/WHO_HSE_EPR_2008_3w.pdf)) (2008).
23. Wood, J. W., Ferrell, R. J. & DeWitte-Aviña, S. N. The temporal dynamics of the fourteenth-century Black Death: new evidence from ecclesiastical records. *Hum. Biol.* **75**, 427–448 (2003).
24. Joosten, M. H. A. J., Cosijnsen, T. J. & De Wit, P. J. G. Host resistance to a fungal tomato pathogen lost by a single base pair change in an avirulence gene. *Nature* **367**, 384–386 (1994).
25. Xu, L. *et al.* Nonlinear effect of climate on plague during the third pandemic in China. *Proc. Natl Acad. Sci. USA* (2011).
26. Keeling, M. K. & Gilligan, C. A. Metapopulation dynamics of bubonic plague. *Nature* **407**, 903–906 (2000).
27. Barrett, R., Kuzawa, C. W., McDade, T. & Armelagos, G. J. Emerging and re-emerging infectious diseases: the third epidemiologic transition. *Annu. Rev. Anthropol.* **27**, 247–271 (1998).
28. Singer, M. & Clair, S. Syndemics and public health: reconceptualizing disease in bio-social context. *Med. Anthropol. Q.* **17**, 423–441 (2003).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank W. White (deceased), J. Bekvalac and R. Redfern from the Museum of London Centre for Human Bioarchaeology for access to samples, M. Kircher and S. Forrest for assistance with computational analysis, G. Wright for support throughout the project, past and present members of the McMaster Ancient DNA Centre for support throughout the project, and D. Poinar for constructive comments on earlier versions of the manuscript. We also thank S. Pääbo and the Max Planck Institute of Evolutionary Anthropology for use of their clean room facilities and molecular biology lab. Funding was provided by the Carl Zeiss Foundation (J.K.), the Human Genetics department of the Medical faculty in Tübingen (J.K.), the Canada Research Chairs program (H.N.P., G.B.G.), the Canadian Institute for Health Research (H.N.P.), the Social Science and Humanities Research Council of Canada (H.N.P.), the Michael G. DeGroote Institute for Infectious Disease Research (H.N.P., B.K.C., D.J.D.E.), an Early Research award from the Ontario Ministry of Research and Education (H.N.P.), the Natural Sciences and Engineering Research Council of Canada (D.J.D.E.), the James S. McDonnell Foundation (D.J.D.E.), and the University at Albany Research Foundation and Center for Social and Demographic Analysis and the Wenner-Gren Foundation (S.N.D.).

**Author Contributions** K.I.B., S.N.D., D.J.D.E., J.K. and H.N.P. conceived the project. K.I.B., S.N.D., S.S. and J.W. performed skeletal sampling. K.I.B., J.K. and V.J.S. carried out laboratory work. H.A.B., K.I.B., J.K., M.M. and H.N.P. designed experiments. K.I.B., G.B.G., J.K., H.N.P., V.J.S. and N.W. analysed the data. B.K.C., D.J.D.E., D.A.H. and J.B.M. provided valuable interpretations. P.B. provided technical support. K.I.B., J.K. and H.N.P. wrote the paper.

**Author Information** Sequencing data have been deposited in GenBank under the accession number SRA045745.1. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). This paper is distributed under the terms of the Creative Commons Attribution Non-Commercial-Share Alike licence, and is freely available to all readers at [www.nature.com/nature](http://www.nature.com/nature). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.K. ([johannes.krause@uni-tuebingen.de](mailto:johannes.krause@uni-tuebingen.de)) or H.N.P. ([poinarh@mcmaster.ca](mailto:poinarh@mcmaster.ca)).

# A natural polymorphism alters odour and DEET sensitivity in an insect odourant receptor

Maurizio Pellegrino<sup>1†</sup>, Nicole Steinbach<sup>1†</sup>, Marcus C. Stensmyr<sup>3</sup>, Bill S. Hansson<sup>3</sup> & Leslie B. Vosshall<sup>1,2</sup>

Blood-feeding insects such as mosquitoes are efficient vectors of human infectious diseases because they are strongly attracted by body heat, carbon dioxide and odours produced by their vertebrate hosts. Insect repellents containing DEET (*N,N*-diethyl-*meta*-toluamide) are highly effective, but the mechanism by which this chemical wards off biting insects remains controversial despite decades of investigation<sup>1–11</sup>. DEET seems to act both at close range as a contact chemorepellent, by affecting insect gustatory receptors<sup>12</sup>, and at long range, by affecting the olfactory system<sup>1–11</sup>. Two opposing mechanisms for the observed behavioural effects of DEET in the gas phase have been proposed: that DEET interferes with the olfactory system to block host odour recognition<sup>1–7</sup> and that DEET actively repels insects by activating olfactory neurons that elicit avoidance behaviour<sup>8–11</sup>. Here we show that DEET functions as a modulator of the odour-gated ion channel formed by the insect odourant receptor complex<sup>13,14</sup>. The functional insect odourant receptor complex consists of a common co-receptor, ORCO (ref. 15) (formerly called OR83B; ref. 16), and one or more variable odourant receptor subunits that confer odour selectivity<sup>17</sup>. DEET acts on this complex to potentiate or inhibit odour-evoked activity or to inhibit odour-evoked suppression of spontaneous activity. This modulation depends on the specific odourant receptor and the concentration and identity of the odour ligand. We identify a single amino-acid polymorphism in the second transmembrane domain of receptor OR59B in a *Drosophila melanogaster* strain from Brazil that renders OR59B insensitive to inhibition by the odour ligand and modulation by DEET. Our data indicate that natural variation can modify the sensitivity of an odour-specific insect odourant receptor to odour ligands and DEET. Furthermore, they support the hypothesis that DEET acts as a molecular ‘confusant’ that scrambles the insect odour code, and provide a compelling explanation for the broad-spectrum efficacy of DEET against multiple insect species.

Previous work has shown that the odour of *Drosophila* food potently attracts adult *D. melanogaster* vinegar flies and that DEET blocks this attraction<sup>5,7</sup>. The behavioural effects of DEET require an intact olfactory system and the olfactory co-receptor ORCO<sup>7</sup>. These results implicated the olfactory system in the observed behavioural effects but failed both to distinguish between the two competing models of action for DEET and to determine whether DEET acts on the odour-specific odourant receptors, ORCO or both. We carried out electrophysiological recordings of *Drosophila* olfactory sensory neurons (OSNs) to test these competing possibilities.

In response to the suggestion that DEET and odours may interact in the vapour phase<sup>9,10</sup>, we first quantified the respective amounts of vapour-phase 1-octen-3-ol emitted from the stimulus pipette in the presence and absence of DEET, using solid-phase microextraction (SPME) followed by gas chromatography mass spectroscopy analysis (GC–MS). The SPME measurements coupled to GC–MS (Fig. 1a) showed that the addition of a second filter paper containing pure

DEET in the stimulus pipette had no significant effect on the release of 1-octen-3-ol ( $10^{-2}$  dilution). Thus, we can rule out any fixative role of DEET under the conditions used here.

We next performed extracellular recordings to measure the effect of DEET on responses elicited by odours in *Drosophila* OSNs housed within the ab2 (Fig. 1a and Supplementary Fig. 1) or ab3 (Supplementary Fig. 2) olfactory hairs, or sensilla, on the fly antenna. Each of these sensilla houses two OSNs expressing different odourant receptors with unique odour response profiles<sup>17</sup>. We measured the activity of these OSNs simultaneously and compared their responses to odour with and without co-presentation of DEET (Fig. 1b, c).

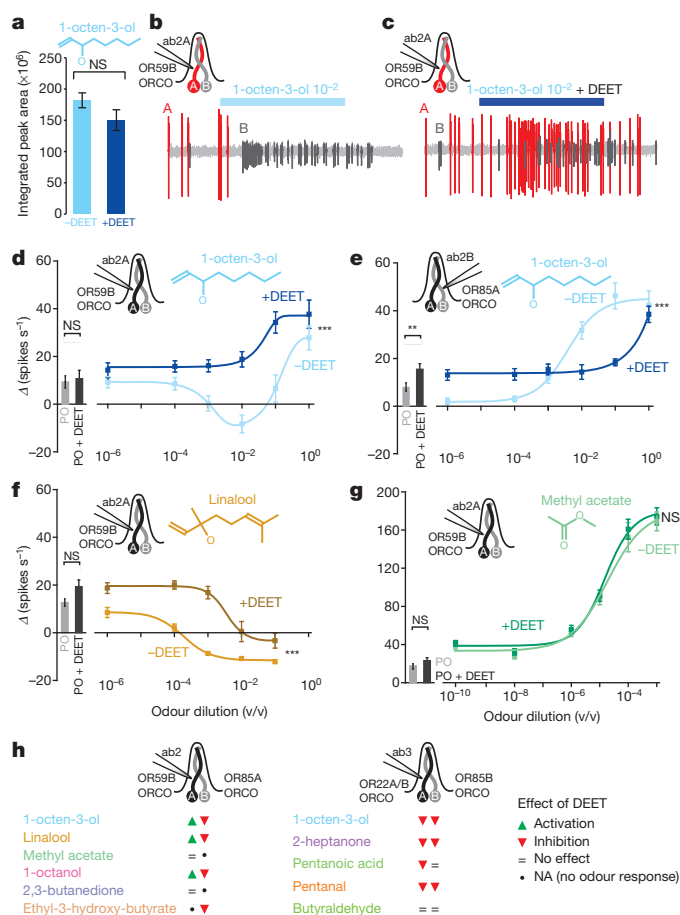
The effect of DEET on four OSNs stimulated with ten structurally diverse odours was complex and dependent on odourant receptor, odour and concentration. In some OSNs, DEET suppressed odour-mediated inhibition (Fig. 1d, f and Supplementary Fig. 1a), in others it decreased odour-induced activation (Fig. 1e, Supplementary Fig. 1b, d, e and Supplementary Fig. 2a–g) and in others it had no effect (Fig. 1g and Supplementary Figs 1c and 2h–j). Moreover, the effects of DEET were strongly concentration dependent, such that high odour concentrations often overcame the effects of DEET (Fig. 1 and Supplementary Figs 1 and 2). DEET presented alone, without odour stimuli, elicited no response above that evoked by solvent in ab2A and ab3A neurons, slightly activated ab2B neurons and slightly inhibited ab3B neurons; but responses were considerably smaller than those elicited by cognate odour ligands (Supplementary Fig. 3). Therefore, DEET alone has a negligible effect on olfactory responses in ab2 and ab3 neurons.

Notably, 1-octen-3-ol presented in a dilution of  $10^{-2}$  had opposite effects on the two neurons housed in ab2 sensilla, inhibiting the ab2A neuron expressing OR59B–ORCO (Fig. 1d) and activating the ab2B neuron expressing OR85A–ORCO (Fig. 1e). Co-application of DEET inverted OSN responses to odour, leading to activation of the ab2A neuron (Fig. 1d) and suppressing the odour-induced activation of the ab2B neuron (Fig. 1e). Similar opposite effects of DEET were observed when the ab2 sensillum was stimulated with a different odour, 1-octanol (Supplementary Fig. 1a, b).

Taken together, our results support the hypothesis that DEET acts as a molecular confusant, scrambling the *Drosophila* odour code by direct modulation of odourant receptor activity dependent on the type of odour and its concentration (Fig. 1h). Recent work examining the effect of DEET on mosquito odourant receptors in heterologous cells supports this hypothesis<sup>18</sup>.

Because the effects of DEET varied with the specific OSN and odour tested, it seems unlikely that DEET acts directly and solely on the conserved co-receptor ORCO, which is co-expressed in all the OSNs examined here. To determine whether DEET acts on the odour-specific odourant receptor subunit, we focused on the pharmacology of the OR59B–ORCO complex in ab2A OSNs. 1-octen-3-ol inhibits basal activity of OR59B–ORCO at low concentrations but acts as an agonist at high concentrations (Fig. 1d). DEET interfered with inhibition of

<sup>1</sup>Laboratory of Neurogenetics and Behaviour, The Rockefeller University, 1230 York Avenue, Box 63, New York, New York 10065, USA. <sup>2</sup>Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, Box 63, New York, New York 10065, USA. <sup>3</sup>Max Planck Institute for Chemical Ecology, Department of Evolutionary Neuroethology, Hans Knöll Strasse 8, 07745 Jena, Germany. †Present addresses: Department of Molecular & Cell Biology, University of California, Berkeley, California 94720 USA (M.P.); Integrated PhD Program in Cellular, Molecular and Biomedical Studies, Columbia University, New York, New York 10032, USA (N.S.).



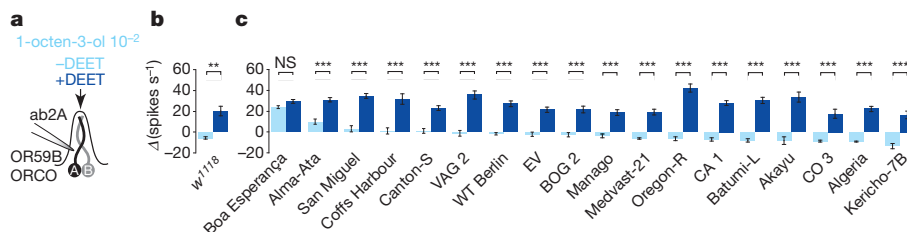
OR59B-ORCO by 1-octen-3-ol, 1-octanol and linalool, but had no effect on odour-dependent activation by methyl acetate and 2,3-butanedione (Fig. 1g and Supplementary Fig. 1c). Notably, DEET had no effect on the OR59B-ORCO activation seen at higher concentrations of 1-octen-3-ol. This selective effect on inhibition might be explained by the presence on the OR59B receptor of distinct 1-octen-3-ol-interaction sites, a high-affinity site that inhibits the odorant receptor complex and is modulated by DEET and a low-affinity DEET-independent site that activates the odorant receptor complex.

To investigate the mechanistic basis of OR59B modulation by DEET, we turned to analysis of this receptor in *D. melanogaster* strains collected around the world. Polymorphisms in natural populations have been previously connected to different sensitivity to odours in

humans<sup>19,20</sup>, and oxygen and carbon dioxide sensing in the nematode *Caenorhabditis elegans*<sup>21</sup>. We reasoned that naturally occurring polymorphisms in insect odorant receptors might modify odorant receptor/odorant interaction sites and affect their sensitivity to DEET. To search for putative polymorphisms that affect DEET responses, we assessed responses of ab2A neurons to 1-octen-3-ol in  $10^{-2}$  dilution in the absence or presence of DEET in 18 wild-type *D. melanogaster* strains from locations around the world, and compared these responses with those obtained in the  $w^{1118}$  laboratory control strain (Fig. 2a, b and Supplementary Fig. 4a). In each strain, ab2 sensilla were identified by the characteristic size and location of the sensilla and responses of the ab2A cell to its cognate ligand, methyl acetate (data not shown). In 17 of the 18 strains, DEET increased responses of ab2A neurons to  $10^{-2}$  1-octen-3-ol (Fig. 2b). However, ab2A neurons in the Brazilian strain Boa Esperança were not inhibited by 1-octen-3-ol at any concentration tested and were therefore insensitive to modulation by DEET (Figs 2c and 3a, b and Supplementary Fig. 4b). In addition to the loss of inhibition by 1-octen-3-ol, the ab2A cell in the Brazilian strain showed robust activation by 1-octanol and ethyl hexanoate, odours that normally inhibit the ab2A cell in wild-type strains. Inhibition by linalool was equivalent in wild-type and Boa Esperança strains (Fig. 3e). Excitatory responses to methyl acetate, ethyl acetate and 2,3-butanedione, both in the absence and presence of DEET, did not differ when compared with the corresponding  $w^{1118}$  neuron (Fig. 3c, d and Supplementary Fig. 5; data not shown). In control experiments, we confirmed that the odour response profiles of ab2A and ab2B OSNs in the Brazilian strain are otherwise similar to that of our  $w^{1118}$  control strain (Fig. 3f and Supplementary Fig. 5).

We proposed that a genetic polymorphism in *Or59b* in the Boa Esperança strain may account for the changed responses to odour and DEET. We therefore sequenced and compared the coding region of *Or59b* in the 19 strains with the published *Or59b* sequence (NCBI reference sequence, NP\_523822.1), and found seven missense polymorphisms and 36 silent polymorphisms among all strains (Supplementary Table 1 and Supplementary Fig. 6). The protein sequence of OR59B in Boa Esperança is referred to as OR59B<sup>Boa</sup> and varies from the NCBI reference at four amino-acid residues (Val41Phe, Val91Ala, Tyr376Ser and Val388Ala). Among these, two are unique to this strain: Val41Phe, located in the amino terminus near transmembrane domain 1 (TM1), and Val91Ala, located within TM2 (Fig. 4a, b and Supplementary Fig. 6). On the basis of our within-strain sampling, we detected only one protein variant per strain except for the  $w^{1118}$  control strain, for which we identified two sequences: one identical to the published OR59B sequence (OR59B<sup>NCBI REF</sup>), and one containing two missense changes (OR59B<sup>M352I T376S</sup>; Fig. 4a and Supplementary Table 1). We analysed electrophysiological recordings obtained from the  $w^{1118}$  control strain for each odour tested and found no evidence that the responses sort into two phenotypically separable clusters. Therefore, we assume that the OR59B<sup>NCBI REF</sup> and OR59B<sup>I352 S376</sup> haplotypes are functionally equivalent, at least for the odours tested in this study. The coding sequences of *Orco* in the  $w^{1118}$  and Boa Esperança strains did not differ from the NCBI reference (data not shown), which suggests that the protein sequence variations in the odour-specific subunit OR59B, rather than the co-receptor ORCO, eliminate inactivation by low concentrations of 1-octen-3-ol and thereby render the odorant receptor complex insensitive to modulation by DEET.

To test the functional consequences of the four OR59B missense changes in the Boa Esperança strain, we generated transgenic flies carrying receptor variants each containing one of the four changes (Val41Phe, Val91Ala, Tyr376Ser or Val388Ala), a combination of the two unique to Boa Esperança (Val41Phe and Val91Ala) or those shared with other strains (Tyr376Ser and Val388Ala), based on the OR59B<sup>NCBI REF</sup> backbone. OR59B variants were selectively expressed in the *Drosophila* *Ahelo* 'empty neuron' system<sup>17,22</sup>, in which the endogenous odour-specific odorant receptors in ab3A OSNs were replaced



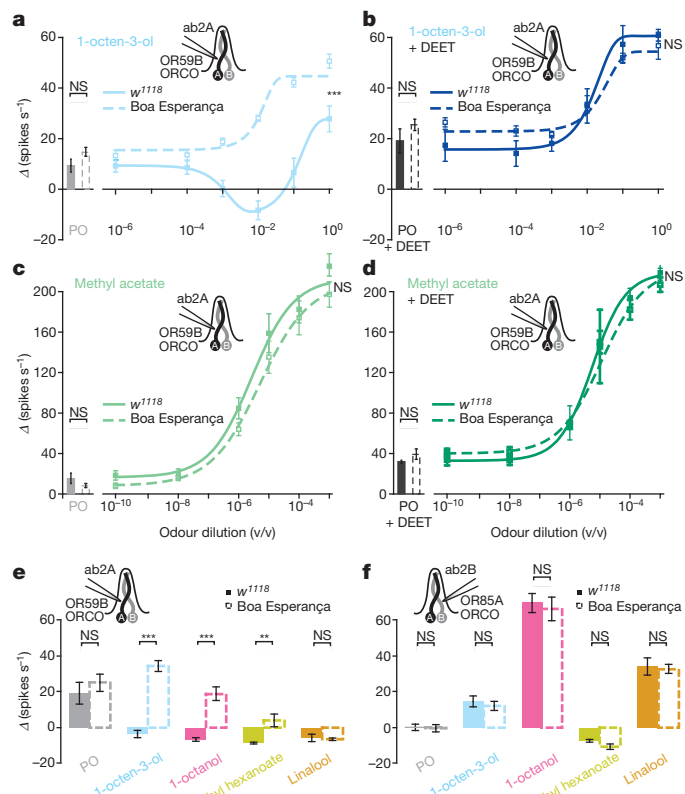
**Figure 2 | OR59B-ORCO sensitivity to DEET varies across wild-type *D. melanogaster* strains.** **a**, Schematic of the screening protocol: 10<sup>-2</sup> 1-octen-3-ol was delivered in the absence and presence of DEET. **b–c**, Bar plots of odour-

evoked responses of the *w*<sup>1118</sup> strain (**b**) and 18 wild-type strains (**c**) to 10<sup>-2</sup> 1-octen-3-ol in the absence (light blue) or presence (dark blue) of DEET (*t*-test with Bonferroni correction; mean ± s.e.m., *n* = 10–17).

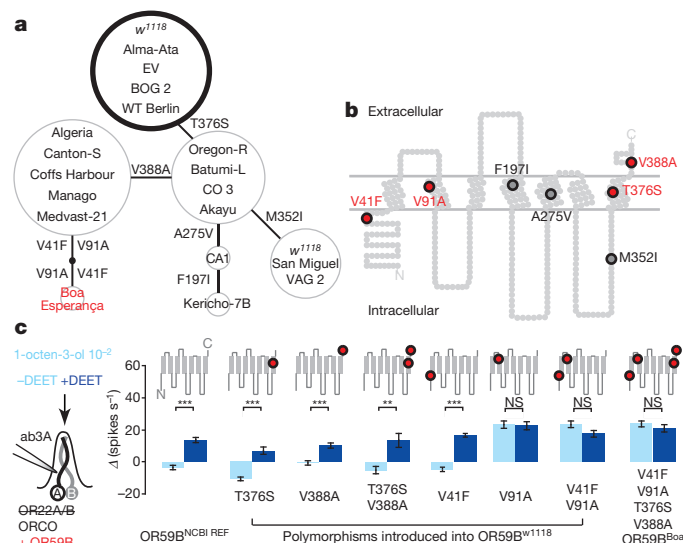
with our OR59B mutants (Fig. 4c and Supplementary Fig. 7). As expected, 10<sup>-2</sup> 1-octen-3-ol caused inhibition of ab3A neurons expressing OR59B<sup>NCBI REF</sup> and activation of ab3A neurons expressing OR59B<sup>Boa</sup> (Fig. 4c). Whereas OR59B<sup>T376A</sup>, OR59B<sup>V388A</sup> and OR59B<sup>T376A V388A</sup> showed normal inhibition to this odour, any variant of OR59B containing the Val91Ala change showed a loss of odour inhibition by 1-octen-3-ol and insensitivity to DEET (Fig. 4c). This demonstrates that the Val91Ala change is sufficient to phenocopy the electrophysiological properties of the endogenous Boa Esperança OR59B (Fig. 4c). It has previously been shown that responses of

OR59B expressed in the empty neuron faithfully recapitulate receptor function measured in the endogenous ab2A neuron<sup>23</sup>. We therefore assume that a strain carrying only the OR59B<sup>V91A</sup> polymorphism would have the same phenotype as Boa Esperança.

DEET shows behavioural efficacy in insects as diverse as *Drosophila*<sup>5,7</sup> and mosquitoes<sup>1–4,6,8–11</sup>. We have shown that a single, naturally occurring polymorphism in an odour-specific odorant receptor can modify receptor interactions with an inhibitory odour and render the receptor insensitive to modulation by DEET. These results provide compelling evidence that DEET interacts directly with an odour-specific odorant receptor. Consistent with this, recent work showed that an odour-specific OR subunit is required for the behavioural effects of DEET on mosquito larvae<sup>11</sup>. Our data imply a complexity in ligand-binding interactions within a single insect odorant receptor complex that bears further investigation. The Val91Ala polymorphism is located in the second predicted transmembrane domain but little is known about which domains of this novel class of odour-gated ion channels contribute to ligand binding or ion channel function<sup>13,14</sup>. A recent study implicated the third predicted transmembrane



**Figure 3 | OR59B-ORCO neurons in the Boa Esperança strain are insensitive to modulation by DEET.** **a–d**, Dose–response curves of the OR59B-ORCO ab2A OSN in wild-type *w*<sup>1118</sup> (solid line) and Boa Esperança (dashed line) strains stimulated with increasing concentrations of 1-octen-3-ol (**a**, **b**) or methyl acetate (**c**, **d**), with (**b**, **d**) or without (**a**, **c**) DEET (*F*-test with Bonferroni correction; mean ± s.e.m., *n* = 5–14). The dose–response curve of *w*<sup>1118</sup> to 1-octen-3-ol in **a** and **b** is reproduced from Fig. 1d for comparison. Bar plots next to the dose–response curves represent responses of the solvent paraffin oil in the absence (grey bar) or presence (black bar) of DEET (*F*-test with Bonferroni correction; mean ± s.e.m., *n* = 5–11). **e**, **f**, Bar plots comparing responses of OR59B-ORCO in ab2A (**e**) and OR85A-ORCO in ab2B (**f**) in *w*<sup>1118</sup> (solid bar) and Boa Esperança (dashed bar) strains to 10<sup>-2</sup> 1-octen-3-ol, 10<sup>-1</sup> 1-octanol, 10<sup>-1</sup> ethyl hexanoate and 10<sup>-1</sup> linalool (*t*-test with Bonferroni correction; mean ± s.e.m., *n* = 9–11).



**Figure 4 | A single natural polymorphism in OR59B confers insensitivity to DEET.** **a**, Haplotype network of OR59B protein variants. Each circle represents a unique OR59B protein variant, its size proportional to the number of strains containing each variant. Connecting lines show the amino-acid substitutions that separate each variant. The bold circle represents the OR59B<sup>NCBI REF</sup> variant with NCBI accession code NP\_523882.1. The Boa Esperança strain is shown in red. **b**, Snake plot of OR59B showing the location of missense polymorphisms. Changes that differentiate Boa Esperança from the NCBI reference are shown in red. **c**, Bar plots show the responses of *Or59b* variants ectopically expressed in ab3A neurons lacking endogenous OR22A and OR22B to 10<sup>-2</sup> 1-octen-3-ol in the absence (light blue) or presence (dark blue) of DEET. The locations of variant amino acids in OR59B are depicted in the cartoon snake plot on top of each set of bar graphs (*t*-test with Bonferroni correction; mean ± s.e.m., *n* = 7–11).

domain of an insect odorant receptor in ligand interactions<sup>24</sup>, and additional structure–function work of this nature will ultimately reveal how these membrane proteins interact with odorants and modulators including DEET. Although Val and Ala are both amino acids with small aliphatic side chains, Val–Ala substitutions have been shown to affect other cation channels<sup>25</sup>. It therefore is plausible that this change would affect the function of the odour-gated ion channel subunit encoded by OR59B. We speculate that the Val91Ala polymorphism inactivates a high-affinity binding site for 1-octen-3-ol that locks the receptor into a closed configuration at low odour concentration. A separate site on the receptor would have a low-affinity binding site that would lead to activation. In this model, DEET would selectively interfere with the high-affinity binding site. Future investigation of the structure–function relation of this receptor is needed to test these ideas. Genetic insensitivity to DEET has previously been shown to exist in both *Drosophila* flies<sup>5</sup> and *Aedes aegypti* mosquitoes<sup>10</sup> but the genes responsible remain unknown. It will be interesting to investigate whether accumulated odorant receptor polymorphisms contribute to these phenotypes.

It has recently been proposed that DEET directly activates behavioural repulsion through the activation of odorant receptors that mediate avoidance behaviours<sup>8–10</sup>. The insect odorant receptor repertoire is highly diverse with very low protein similarity across insect species<sup>26–28</sup>. Furthermore, different species respond very selectively to host odour cues that meet disparate ecological needs<sup>29,30</sup>. It seems unlikely that a single molecule like DEET would activate a different yet similarly potent repulsive behaviour in all insects tested. Instead, our data support the hypothesis that DEET is a broad-selectivity insect odorant receptor modulator that alters the fine-tuning of the insect olfactory system. DEET-mediated scrambling of the odour code would interfere with behavioural responses as diverse as mosquitoes orienting to host odours produced by humans<sup>29</sup> and the attraction of *Drosophila* to yeast on rotting fruit<sup>30</sup>.

## METHODS SUMMARY

**Fly strains and molecular biology.** *D. melanogaster* stocks were maintained on conventional cornmeal–agar–molasses medium in a 12-h-light, 12-h-dark cycle at 25 °C. Details of molecular biology manipulations, all primers and fly strains are in Methods.

**Single-sensillum extracellular recordings.** Recordings of female fly antennae were performed as described previously<sup>7</sup> and are detailed in Methods. The respective amounts of 1-octen-3-ol emitted from the stimulus pipettes with and without DEET was investigated through SPME and linked GC–MS analysis as detailed in Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 17 September 2010; accepted 8 August 2011.

Published online 21 September 2011.

1. Davis, E. E. & Sokolove, P. G. Lactic acid-sensitive receptors on the antennae of the mosquito, *Aedes aegypti*. *J. Comp. Physiol. A* **105**, 43–54 (1976).
2. McIver, S. B. A model for the mechanism of action of the repellent DEET on *Aedes Aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **18**, 357–361 (1981).
3. Dogan, E. B., Ayres, J. W. & Rossignol, P. A. Behavioural mode of action of DEET: inhibition of lactic acid attraction. *Med. Vet. Entomol.* **13**, 97–100 (1999).
4. Dogan, E. B. & Rossignol, P. A. An olfactometer for discriminating between attraction, inhibition, and repellency in mosquitoes (Diptera: Culicidae). *J. Med. Entomol.* **36**, 788–793 (1999).
5. Reeder, N. L., Ganz, P. J., Carlson, J. R. & Saunders, C. W. Isolation of a deet-insensitive mutant of *Drosophila melanogaster* (Diptera: Drosophilidae). *J. Econ. Entomol.* **94**, 1584–1588 (2001).
6. Kline, D. L., Bernier, U. R., Posey, K. H. & Barnard, D. R. Olfactometric evaluation of spatial repellents for *Aedes aegypti*. *J. Med. Entomol.* **40**, 463–467 (2003).

7. Ditzgen, M., Pellegrino, M. & Vosshall, L. B. Insect odorant receptors are molecular targets of the insect repellent DEET. *Science* **319**, 1838–1842 (2008).
8. Xia, Y. *et al.* The molecular and cellular basis of olfactory-driven behavior in *Anopheles gambiae* larvae. *Proc. Natl Acad. Sci. USA* **105**, 6433–6438 (2008).
9. Syed, Z. & Leal, W. S. Mosquitoes smell and avoid the insect repellent DEET. *Proc. Natl Acad. Sci. USA* **105**, 13598–13603 (2008).
10. Stanczyk, N. M., Brookfield, J. F., Ignell, R., Logan, J. G. & Field, L. M. Behavioral insensitivity to DEET in *Aedes aegypti* is a genetically determined trait residing in changes in sensillum function. *Proc. Natl Acad. Sci. USA* **107**, 8575–8580 (2010).
11. Liu, C. *et al.* Distinct olfactory signaling mechanisms in the malaria vector mosquito *Anopheles gambiae*. *PLoS Biol.* **8**, e1000467 (2010).
12. Lee, Y. S., Kim, S. H. & Montell, C. Avoiding DEET through insect gustatory receptors. *Neuron* **67**, 555–561 (2010).
13. Sato, K. *et al.* Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–1006 (2008).
14. Wicher, D. *et al.* *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**, 1007–1011 (2008).
15. Larsson, M. C. *et al.* Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* **43**, 703–714 (2004).
16. Vosshall, L. B. & Hansson, B. S. A unified nomenclature system for the insect olfactory co-receptor. *Chem. Senses* **36**, 497–498 (2011).
17. Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
18. Bohbot, J. D. & Dickens, J. C. Insect repellents: modulators of mosquito odorant receptor activity. *PLoS ONE* **5**, e12138 (2010).
19. Keller, A., Zhuang, H., Chi, Q., Vosshall, L. B. & Matsunami, H. Genetic variation in a human odorant receptor alters odour perception. *Nature* **449**, 468–472 (2007).
20. Menashe, I. *et al.* Genetic elucidation of human hyperosmia to isovaleric acid. *PLoS Biol.* **5**, e284 (2007).
21. McGrath, P. T. *et al.* Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* **61**, 692–699 (2009).
22. Gross, S. P., Guo, Y., Martinez, J. E. & Welte, M. A. A determinant for directionality of organelle transport in *Drosophila* embryos. *Curr. Biol.* **13**, 1660–1668 (2003).
23. Hallem, E. A., Ho, M. G. & Carlson, J. R. The molecular basis of odor coding in the *Drosophila* antenna. *Cell* **117**, 965–979 (2004).
24. Nichols, A. S. & Luetje, C. W. Transmembrane segment 3 of *Drosophila melanogaster* odorant receptor subunit 85b contributes to ligand-receptor interactions. *J. Biol. Chem.* **285**, 11854–11862 (2010).
25. Laish-Farkash, A. *et al.* A novel mutation in the HCN4 gene causes symptomatic sinus bradycardia in Moroccan Jews. *J. Cardiovasc. Electrophysiol.* **21**, 1365–1372 (2010).
26. Robertson, H. M., Warr, C. G. & Carlson, J. R. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **100** (suppl. 2), 14537–14542 (2003).
27. Robertson, H. M. & Wanner, K. W. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res.* **16**, 1395–1403 (2006).
28. Bohbot, J. *et al.* Molecular characterization of the *Aedes aegypti* odorant receptor gene family. *Insect Mol. Biol.* **16**, 525–537 (2007).
29. Takken, W. & Knols, B. G. Odor-mediated behavior of Afrotropical malaria mosquitoes. *Annu. Rev. Entomol.* **44**, 131–157 (1999).
30. Stensmyr, M. C., Giordano, E., Balloi, A., Angioy, A. M. & Hansson, B. S. Novel natural ligands for *Drosophila* olfactory receptor neurons. *J. Exp. Biol.* **206**, 715–724 (2003).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank C. Bargmann, K. Lee, K. Scott, L. Stowers and members of the Vosshall lab for discussion and comments on the manuscript; and K. Weniger for technical assistance with the SPME and GC–MS experiments. This work was funded in part by a grant to R. Axel and L.B.V. from the Foundation for the National Institutes of Health through the Grand Challenges in Global Health Initiative and by a grant to L.B.V. from the NIH (R01 DC008600). L.B.V. is an investigator of the Howard Hughes Medical Institute. M.C.S. and B.S.H. are supported by the Max Planck Society.

**Author Contributions** M.P. carried out all the experiments and analysed the data. N.S. contributed to sequencing *Or59b* in the 19 strains and generated the *Or59b* mutants. M.C.S. and B.S.H. designed and supervised the SPME collections and GC–MS analysis in Fig. 1a. M.P. and L.B.V. together designed the experiments, interpreted the results, produced the figures and wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to L.B.V. ([leslie.vosshall@rockefeller.edu](mailto:leslie.vosshall@rockefeller.edu)).

## METHODS

**Genomic DNA.** DNA was prepared according to the Quick Fly Genomic DNA Prep protocol from the Berkeley Drosophila Genome Project (<http://www.fruitfly.org/about/methods/inverse.pcr.html>). DNA (1.5 µl) was used for amplification using the KOD PCR Kit (Novagen). For *Or59b*, primers were designed to anneal to the 5' and 3' untranslated regions of the *w<sup>1118</sup> Or59b* locus: 5'-gaattcTCCGGG TATAAAGTGCAGGTGCTGGCACCG-3' (forward); 5'-ctcgagGCTCTTTTTCGGGGGGCTCATGGGTGCAG-3' (reverse).

*Orco* was amplified using primers that amplify the complete coding region: 5'-gaattcATGACAACCTCGATGCAG-3' (forward); 5'-caattgCTTGAGCTGCA CCAGCACCA-3' (reverse).

PCR products were cloned into pGEM-T Easy (Promega Corporation), sequenced (GENEWIZ, Inc.) and analysed using SeqMan software (DNASTAR, Inc.). For each strain, at least four independent samples were analysed, derived from at least two different genomic preparations and two different PCR reactions. These were sequenced and compared to NCBI reference sequences for each gene (*Or59b*: NM\_079098.1; *Orco*: NM\_079511.4).

**Complementary DNA preparation and transgenic flies.** Total RNA was extracted from *w<sup>1118</sup>* and Boa Esperança antennae using the RNeasy Mini Kit (QIAGEN).

Complementary DNA (cDNA) synthesis was performed according to the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) using oligo(dT) primers. *Or59b* cDNA from both *w<sup>1118</sup>* and Boa Esperança was amplified using these gene-specific primers: 5'-gaattcATGGCGGTGTCAAGCT AATCAAACCG-3' (forward); 5'-ctcgagTTACTGGAAGTCTCGGCCAGATT CA-3' (reverse).

PCR products representing full-length *w<sup>1118</sup> Or59b<sup>NCBI REF</sup>* and *Or59b<sup>Boa</sup>* cDNAs were cloned into pGEM-T Easy, completely sequenced and subcloned into the pUAST attB vector<sup>31</sup> using EcoRI and XhoI restriction sites.

Single point mutations were introduced into the *w<sup>1118</sup> Or59b<sup>NCBI REF</sup>* cDNA by directed PCR mutagenesis. Two independent reactions were prepared: one contained the forward primer with the desired mutation and the reverse SP6 vector primer (5'-ATTTAGGTGACACTATAG-3'). The second contained the reverse mutating primer and the forward T7 vector primer (5'-TAATACGACTCAC TATAGGG-3'). PCR products from the reactions were purified and 1 µl of each was used as a template and mixed in a second round of amplification with T7 and SP6 primers to obtain the full gene. For each mutagenesis, the final PCR product was purified and subcloned in pGEM-T Easy, and the complete *Or59b* cDNA carrying the induced mutations was sequenced for verification and compared with the *Or59b<sup>NCBI REF</sup>* sequence.

The double mutants *Or59b<sup>V41F V91A</sup>* and *Or59b<sup>T376S V388A</sup>* were generated using *Or59b<sup>V41F</sup>* or *Or59b<sup>T376S</sup>* as a template and a second round of mutagenesis was implemented with the corresponding primers.

The following primers were used. *Or59b<sup>V41F</sup>*: 5'-CCGCCGAAGGAGGGATT CCTGCGCTACGTGT-3' (forward); 5'-ACACGTAGCGCAGGAATCCCTCC TTCGGCGG-3' (reverse). *Or59b<sup>V91A</sup>*: 5'-AGGTGTGCATCAATCGGTATGGC GCCTCGG-3' (forward); 5'-CCGAGGCGCCATACGCATTGATGCACACCTC-3' (reverse). *Or59b<sup>T376S</sup>*: 5'-TGAACAGCAACATAAGCGTGCCCAAGTTC GC-3' (forward); 5'-GCGAAGCTGGCCAGCTTATGTTGCTGTTC-3' (reverse). *Or59b<sup>V388A</sup>*: 5'-GCATCATTACAATAGCGGACAAATGAATCT-3' (forward); 5'-AGATTCAATTGTGCGCTATTGTAATGATGC-3' (reverse).

Transgenic animals were generated in the *w<sup>1118</sup>* genetic background (Genetic Services, Inc.) using the phiC31-based integration system<sup>31</sup> targeted at the attP2-docking site on chromosome II (ref. 32).

**Fly stocks.** *Drosophila melanogaster* stocks were maintained on conventional cornmeal–agar–molasses medium in a 12-h-light, 12-h-dark cycle at 25 °C. The *w<sup>1118</sup>* strain was used as wild-type control.

The following wild-type strains were used: Akayu [*Drosophila* Genetic Resource Center (DGRC) #103389; origin, Japan]; Algeria (isogenic for II and III chromosomes, DGRC #103390; origin, Algeria); Alma-Ata (DGRC #103391; origin, Kazakhstan); Canton-S (isogenic for II and III, lab stock; origin, Ohio, USA); CA1 (Bloomington *Drosophila* Stock Center #3846; origin, Cape Town, South Africa); Coffs Harbour (DGRC #103411; origin, New South Wales, Australia); Kericho-7B (DGRC #103428; origin, Kericho, Kenya); Manago (isogenic for II and III, DGRC #103433; origin, Hawaii, USA); Oregon-R (isogenic for II and III, lab stock; origin, Oregon, USA); San Miguel (isogenic for II and III, DGRC #103450; origin, Buenos Aires, Argentina); WT Berlin (isogenic for II and III, Heisenberg laboratory, Würzburg, Germany; origin, Berlin, Germany); Batumi-L (DGRC #103396; origin, Batumi, Georgia); Boa Esperança (DGRC #103400; origin, Minas Gerais, Brazil); BOG2 (Bloomington #3842; origin, Bogota, Colombia); CO3 (Bloomington #3848; origin, Commack, New York, USA); EV (Bloomington #3851; origin, Ellenville, New York, USA); Medvast-21 (DGRC #103435; origin, Finland); VAG 2 (Bloomington #3876; origin, Athens, Greece).

The following mutant alleles and transgenic flies were used: *Or22a/b<sup>Ahalo</sup>* (ref. 33) and *Or22a-Gal4* (ref. 34). The genotypes of the flies used for Fig. 4c and Supplementary Fig. 8 are as follows: *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b* (labelled *Or59b<sup>NCBI REF</sup>* in the figure), *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>V41F</sup>* (*V41F*), *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>V91A</sup>* (*V91A*), *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>V41F V91A</sup>* (*V41F V91A*), *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>T376S</sup>* (*T376S*), *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>V388A</sup>* (*V388A*), *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>T376S V388A</sup>* (*T376S V388A*) and *Or22a/b<sup>Ahalo</sup>, Or22a-Gal4/UAS-Or59b<sup>V41F V91A T376S V388A</sup>* (*V41F V91A T376S V388A*).

**SPME quantification of emitted volatiles.** The effect of DEET on the amount of 1-octen-3-ol emitted from the stimulus pipettes was investigated through SPME and linked GC-MS analysis. Stimulus pipettes, prepared as per the electrophysiology experiments, were loaded either with one filter strip impregnated with 5 µl of 1-octen-3-ol ( $10^{-2}$ ) and with a second strip containing 5 µl of paraffin oil, or with the second strip impregnated with 5 µl of pure DEET. The pipettes were connected to a stimulus controller (Syntech CS 55; [www.syntech.nl](http://www.syntech.nl)) and volatiles emitted from the pipettes during ten puffs, of 2-s duration each, delivered with 1-s intervals, were trapped on a SPME fibre (Supelco blue fibre; 57310-U; polydimethylsiloxane/divinylbenzene, 65-µm coating; <http://www.sigmaaldrich.com>), inserted 2 cm into the pipette tip. After completion of the stimulus cycle, the SPME fibres were immediately retracted and injected into a GC-MS device for quantification. This device (Agilent GC6890N fitted with MS5975B unit; <http://www.agilent.com>) was equipped with a HP5-MS column (Agilent Technologies) and operated as follows. The inlet temperature was set to 250 °C. Desorption time was 1 min. The temperature of the gas chromatography oven was held at 70 °C for 2 min and then increased by 20 °C min<sup>-1</sup> to 280 °C, with the final temperature held for 2 min. For mass spectroscopy, the transfer line was held at 280 °C, the source at 230 °C and the quad at 150 °C. Mass spectra were taken in EI mode (at 70 eV) in the range from 33 m/z to 350 m/z, with a scanning rate of 4.42 scans per second. GC-MS data were processed with the MDS-CHEMSTATION software (Agilent Technologies), and peak areas were autointegrated. Five replicates were collected for each condition and data were plotted as mean  $\pm$  s.e.m. Statistical significance was assessed using a *t*-test.

**Electrophysiology and odors.** Female transgenic flies were recorded at 5 d after adult eclosion. All other flies were recorded at 5–10 d after adult eclosion. Single-sensillum recordings were performed as described previously<sup>35,36</sup>. For each experiment in which we recorded OR59B variants expressed in the ab3A neuron, we verified that responses of endogenous OR59B in the native ab2A neuron showed normal inhibition by  $10^{-2}$  1-octen-3-ol (data not shown). Odorants were obtained from Sigma-Aldrich at high purity and diluted (v/v) in paraffin oil as indicated. DEET was obtained from Alfa Aesar and was applied undiluted. Chemical Abstracts Service (CAS) numbers are as follows: paraffin oil (8012-95-1); 1-octen-3-ol (3391-86-4); pentanal (110-62-3); pentanoic acid (109-52-4); 2-heptanone (110-43-0); 1-octanol (111-87-5); (–)-linalool (126-91-0); methyl acetate (79-20-9); 2,3-butanedione (431-03-8); ethyl hexanoate (123-66-0); butyraldehyde (123-72-8); ethyl-3-hydroxybutyrate (5405-41-4); ethyl acetate (141-78-6); hexanol (111-27-3); DEET (134-62-3).

The desired odour dilution (30 µl) was pipetted onto a filter paper strip (3 mm  $\times$  50 mm) and 30 µl of undiluted DEET or paraffin oil solvent was pipetted onto a second filter paper strip. Both filter paper strips were then carefully inserted into a glass Pasteur pipette. Before any recordings, charcoal-filtered air was forced through the pipette for 1–3 s to remove dead space in the odour delivery system. For actual recordings, charcoal-filtered air was continuously applied to the insect antenna, with odour delivered through the pipette to the fly antennae for 1 s. Each pipette was used at most three times and no more than three sensilla were tested per animal. Sensilla types were identified by size, location on the antenna and responsiveness to known preferred odorants<sup>37</sup>.

Data were collected using AUTOSPIKE (Syntech) and analysed by custom spike-sorting algorithms<sup>35</sup>. Responses were initially classified as excitatory or inhibitory by visual inspection of the responses after odour application. An odour was classified as excitatory if it increased the spontaneous firing rate and inhibitory if it decreased the spontaneous firing rate. The data were then analysed by subtracting average spontaneous activity (expressed as spikes per second) in the 15 s before odour application from activity either in the first 600 ms after odour delivery, for excitatory odorants, or in the first 1 s, for inhibitory odorants. This value is referred to as *A*, and will typically have a negative value for inhibitory odorants and a positive value for excitatory odorants. The onset of odour-evoked responses varied owing to slight variations in the position of the odour delivery system relative to the sensillum being recorded. To correct for this, we calibrated the inferred odour onset on the basis of excitatory responses elicited by control stimuli applied at the beginning of each trial (ab2,  $10^{-5}$  methyl acetate; ab3,  $10^{-5}$  2-heptanone).

**Statistical analysis.** Dose–response curves were fitted with ORIGINPRO 8 (OriginLab) using a logistic function, except for responses to 1-octen-3-ol in Fig. 1d, which used a biphasic function.

Comparisons of paired dose–response curves in Figs 1 and 3 and Supplementary Figs 1, 2 and 4 used an *F*-test to assess the statistical significance of differences between the two curve fits. A two-tailed *t*-test was performed for all comparisons in Fig. 1i (non-paired), Figs 2–4 and Supplementary Figs 3, 4 and 7 (paired). Type I errors were addressed by using a Bonferroni correction for multiple comparisons applied to each set of experiments. Data in Supplementary Fig. 6 were fitted using a linear regression analysis.

The OR59B snake plots in Fig. 4 and Supplementary Fig. 7 were hand-composed on the basis of transmembrane domain predictions generated with the PredictProtein algorithm<sup>38</sup>.

31. Bischof, J., Maeda, R. K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl Acad. Sci. USA* **104**, 3312–3317 (2007).
32. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genet.* **40**, 476–483 (2008).
33. Dobritsa, A. A., van der Goes van Naters, W., Warr, C. G., Steinbrecht, R. A. & Carlson, J. R. Integrating the molecular and cellular basis of odor coding in the *Drosophila* antenna. *Neuron* **37**, 827–841 (2003).
34. Fishilevich, E. & Vosshall, L. B. Genetic and functional subdivision of the *Drosophila* antennal lobe. *Curr. Biol.* **15**, 1548–1553 (2005).
35. Ditzen, M., Pellegrino, M. & Vosshall, L. B. Insect odorant receptors are molecular targets of the insect repellent DEET. *Science* **319**, 1838–1842 (2008).
36. Pellegrino, M., Nakagawa, T. & Vosshall, L. B. Single sensillum recordings in the insects *Drosophila melanogaster* and *Anopheles gambiae*. *J. Vis. Exp.* **36**, 1–5 (2010).
37. Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
38. Rost, B., Yachdav, G. & Liu, J. The PredictProtein server. *Nucleic Acids Res.* **32**, W321–W326 (2004).

# STING is a direct innate immune sensor of cyclic di-GMP

Dara L. Burdette<sup>1</sup>, Kathryn M. Monroe<sup>1</sup>, Katia Sotelo-Troha<sup>1</sup>, Jeff S. Iwig<sup>1,2</sup>, Barbara Eckert<sup>1</sup>, Mamoru Hyodo<sup>3</sup>, Yoshihiro Hayakawa<sup>4</sup> & Russell E. Vance<sup>1</sup>

The innate immune system detects infection by using germline-encoded receptors that are specific for conserved microbial molecules. The recognition of microbial ligands leads to the production of cytokines, such as type I interferons (IFNs), that are essential for successful pathogen elimination. Cytosolic detection of pathogen-derived DNA is one major mechanism of inducing IFN production<sup>1,2</sup>, and this process requires signalling through TANK binding kinase 1 (TBK1) and its downstream transcription factor, IFN-regulatory factor 3 (IRF3). In addition, a transmembrane protein called STING (stimulator of IFN genes; also known as MITA, ERIS, MPYS and TMEM173) functions as an essential signalling adaptor, linking the cytosolic detection of DNA to the TBK1–IRF3 signalling axis<sup>3–7</sup>. Recently, unique nucleic acids called cyclic dinucleotides, which function as conserved signalling molecules in bacteria<sup>8</sup>, have also been shown to induce a STING-dependent type I IFN response<sup>9–12</sup>. However, a mammalian sensor of cyclic dinucleotides has not been identified. Here we report evidence that STING itself is an innate immune sensor of cyclic dinucleotides. We demonstrate that STING binds directly to radiolabelled cyclic diguanylate monophosphate (c-di-GMP), and we show that unlabelled cyclic dinucleotides, but not other nucleotides or nucleic acids, compete with c-di-GMP for binding to STING. Furthermore, we identify mutations in STING that selectively affect the response to cyclic dinucleotides without affecting the response to DNA. Thus, STING seems to function as a direct sensor of cyclic dinucleotides, in addition to its established role as a signalling adaptor in the IFN response to cytosolic DNA. Cyclic dinucleotides have shown promise as novel vaccine adjuvants and immunotherapeutics<sup>9,13</sup>, and our results provide insight into the mechanism by which cyclic dinucleotides are sensed by the innate immune system.

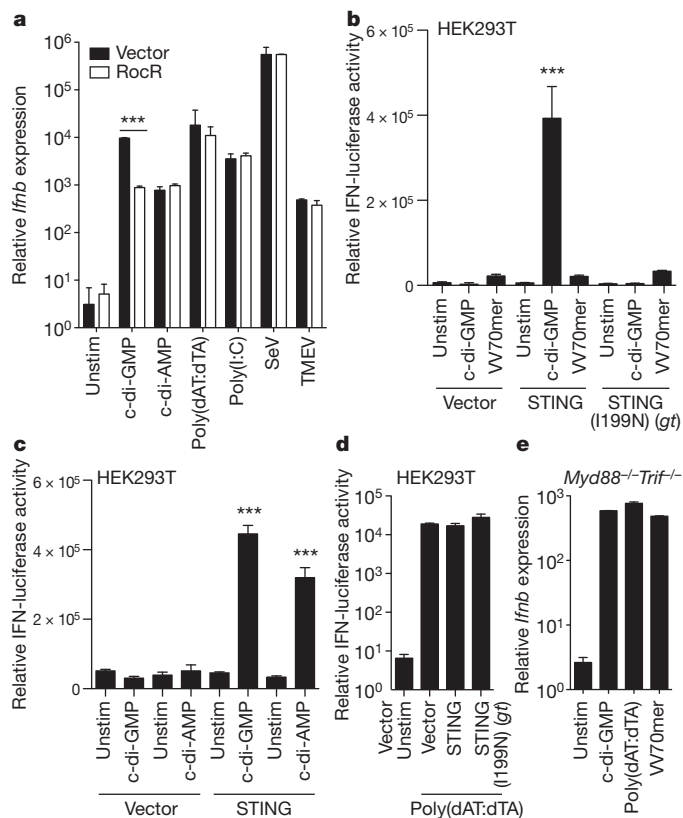
Nucleotides are crucial signalling molecules in all domains of life, but cyclic dinucleotides seem to be produced solely by Bacteria and Archaea. For example, c-di-GMP is a ubiquitous second messenger that regulates biofilm formation, motility and virulence in a diverse range of bacterial species<sup>8</sup>. Recently, cyclic diadenylate monophosphate (c-di-AMP) was discovered to be a bacterial regulatory molecule<sup>14</sup>, although its role remains to be fully characterized. Because they are unique to microorganisms, cyclic dinucleotides are appropriate targets for immune recognition<sup>15</sup>. Indeed, the induction of IFN production by *Listeria monocytogenes* depends on bacterial secretion of cyclic-di-AMP<sup>12</sup>. However, it remains unclear how cyclic dinucleotides are sensed in mammalian cells.

To address the mechanism by which mammalian cells sense cyclic dinucleotides, we first confirmed that cyclic dinucleotides are detected in the host cell cytosol<sup>10</sup> by expressing RocR, a c-di-GMP-specific phosphodiesterase from *Pseudomonas aeruginosa*, in the cytosol of macrophages. In these cells, IFN induction in response to c-di-GMP (but not other stimuli) is tenfold lower than that in control, vector-transduced, cells (Fig. 1a), confirming that the cytosolic presence of c-di-GMP is important for inducing IFN.

To identify candidate cyclic dinucleotide sensors, we sought to identify molecules that could reconstitute the IFN response to cyclic dinucleotides in HEK293T cells, which do not respond to c-di-GMP<sup>10</sup>. Because STING is essential for the IFN response to cyclic dinucleotides<sup>11</sup> and because STING expression is low or undetectable in HEK293T cells (Supplementary Fig. 1 and data not shown), we first expressed STING in HEK293T cells. The overexpression of STING spontaneously induces an IFN reporter<sup>3,6,7</sup>, so we transfected a small amount of *Sting*-encoding vector that by itself was insufficient to induce IFN. Low levels of STING protein were sufficient to reconstitute the responsiveness of HEK293T cells to c-di-GMP (Fig. 1b) and c-di-AMP (Fig. 1c). By contrast, the non-functional goldenticket (*gt*) allele of *Sting* (which results in a STING protein in which asparagine has been substituted for isoleucine, I199N)<sup>11</sup> did not restore responsiveness to c-di-GMP (Fig. 1b). Interestingly, the expression of wild-type STING did not reconstitute the responsiveness of HEK293T cells to double-stranded DNA (dsDNA) oligonucleotides (for example, a 70-base-pair oligonucleotide from vaccinia virus (VV70mer) or IFN-stimulatory DNA (ISD)) that had previously been shown to induce type I IFNs in macrophages through STING<sup>4,16</sup> (Fig. 1b and Supplementary Fig. 2a). By contrast, the induction of IFN by poly(dA:dT)·poly(dT:dA) DNA (denoted poly(dAT:dTA)) was identical in cells that were transfected with wild-type *Sting* and those transfected with *gt*, demonstrating that the RNA polymerase III DNA-sensing pathway<sup>17,18</sup> is intact in these cells and is not responsible for the detection of c-di-GMP (Fig. 1d). As a positive control, *Myd88*<sup>−/−</sup>*Trif*<sup>−/−</sup> immortalized macrophages, which express STING, responded similarly to c-di-GMP, poly(dAT:dTA), VV70mer and ISD (Fig. 1e and Supplementary Fig. 2b). Together, our results show that STING expression is sufficient to restore the responsiveness of HEK293T cells to cyclic dinucleotides but not to DNA.

We next tested whether STING, or perhaps another protein in HEK293T cells, binds to c-di-GMP. We used HEK293T cell lysates in an *in vitro* ultraviolet radiation crosslinking assay to identify putative sensor proteins that interact directly with radiolabelled c-di-GMP (c-di-[<sup>32</sup>P]GMP). We expected to identify directly interacting proteins because only molecules within bond-length proximity are efficiently crosslinked by ultraviolet radiation<sup>19</sup>. We detected a prominent ~40-kDa radiolabelled protein, which corresponds to the predicted molecular weight of monomeric STING, in the lysates of cells transfected with a vector encoding haemagglutinin (HA)-tagged STING (STING–HA) but not in the lysates of cells transfected with a vector encoding STING(I199N)–HA or vector only (Fig. 2a). The ~40-kDa band did not appear when the same lysates were crosslinked in the presence of [<sup>32</sup>P]GTP, implying that crosslinking to c-di-[<sup>32</sup>P]GMP was specific (Fig. 2a). We also observed an ~80-kDa species, which might correspond to a previously reported STING dimer<sup>6</sup> (Fig. 2b). To test the hypothesis that STING crosslinks with c-di-[<sup>32</sup>P]GMP, we immunoprecipitated STING from transfected HEK293T cells and performed the c-di-[<sup>32</sup>P]GMP crosslinking assay on the

<sup>1</sup>Department of Molecular & Cell Biology, University of California, Berkeley, California 94720, USA. <sup>2</sup>Department of Chemistry, Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA. <sup>3</sup>Faculty of Pharmaceutical Science, Hokkaido University, Kita 12, Nishi 6, Kita-ku, Sapporo, Hokkaido 060-0812, Japan. <sup>4</sup>Department of Applied Chemistry, Faculty of Engineering, Aichi Institute of Technology, 1247 Yachigusa, Yakusa, Toyota 470-0392, Japan.



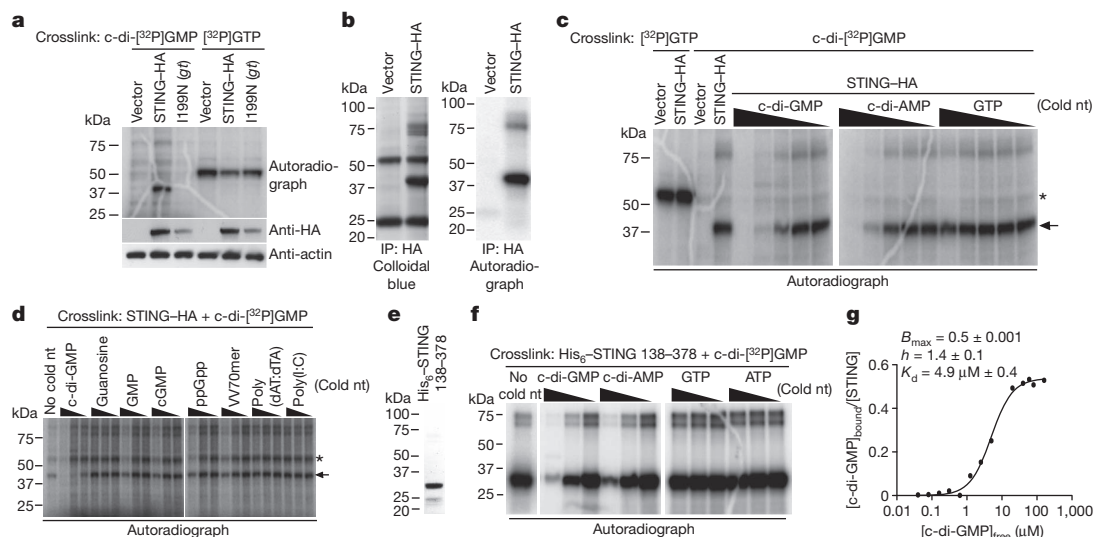
immunoprecipitates. Bands corresponding to the molecular weight of the STING monomer and dimer were identified only in immunoprecipitates from lysates overexpressing STING and not in mock immunoprecipitates from lysates of vector-only-transfected cells (Fig. 2b). Thus, STING seems to bind to c-di-GMP.

To confirm that the binding of c-di-[ $^{32}$ P]GMP to STING is specific, we performed the c-di-GMP crosslinking assay in the presence of

**Figure 1** | STING is sufficient to restore responsiveness to cyclic dinucleotides. **a**, Immortalized *Myd88*<sup>-/-</sup>*Trif*<sup>-/-</sup> macrophages were transduced with retrovirus expressing *rocR* and then stimulated for 6 h. IFN induction was measured by quantitative PCR with reverse transcription (quantitative RT-PCR) and normalized to ribosomal protein 17 (*Rps17*) expression levels. **b–d**, HEK293T cells were transfected as indicated, together with an IFN-luciferase reporter, and luciferase activity was measured 6 h after stimulation. **e**, *Myd88*<sup>-/-</sup>*Trif*<sup>-/-</sup> macrophages were stimulated for 6 h, and IFN induction was measured as in **a**. **a–e**, Data are presented as mean  $\pm$  s.d. ( $n = 3$ ) and are representative of at least three independent experiments. \*\*\*,  $P < 0.001$ . SeV, Sendai virus; TMEV, Theiler's murine encephalomyelitis virus; unstim, unstimulated; VV70mer, a stimulatory dsDNA oligonucleotide derived from vaccinia virus.

unlabelled nucleotides. Unlabelled c-di-GMP and c-di-AMP specifically competed with c-di-[ $^{32}$ P]GMP for binding to STING (Fig. 2c, d). By contrast, GTP, other guanosine derivatives and nucleic acids (including dsDNA) competed away nonspecific binding (Fig. 2c, d, asterisk); however, under our specific assay conditions, these molecules could not compete efficiently with c-di-[ $^{32}$ P]GMP for binding to STING (Fig. 2c, d, arrow). Because the cell cytosol contains high concentrations of GTP (0.1–1 mM), a putative c-di-GMP sensor must have a high degree of specificity for c-di-GMP over GTP. We found that c-di-GMP efficiently crosslinked to STING even in the presence of 1 mM GTP (Fig. 2c).

Although these data imply that STING directly and specifically binds to cyclic dinucleotides, they do not address whether other host proteins might also be required. STING is predicted to encode an amino-terminal domain with multiple transmembrane segments, followed by a globular carboxy-terminal domain (CTD). Because the CTD contains the amino acid substitution that abolishes STING function in *gt* mice (I199N)<sup>11</sup>, we suspected that the CTD might be involved in binding to cyclic dinucleotides. Thus, we subjected purified recombinant His<sub>6</sub>-tagged STING CTD (amino acids 138–378) (Fig. 2e) to the c-di-[ $^{32}$ P]GMP crosslinking assay. We found that the recombinant CTD of STING bound to c-di-[ $^{32}$ P]GMP and that this binding was specifically competed away with cold (unlabelled) c-di-GMP or c-di-AMP but not with cold GTP or ATP (Fig. 2f). We used equilibrium



**Figure 2** | STING binds cyclic dinucleotides. **a**, HEK293T cells were transfected as indicated, and the cell lysates were subjected to an *in vitro* ultraviolet radiation crosslinking assay with c-di-[ $^{32}$ P]GMP. Samples were separated by SDS-PAGE and visualized by autoradiography or western blotting. **b**, HEK293T cells were transfected as in **a**, and anti-HA immunoprecipitates were stained with colloidal blue or subjected to the crosslinking assay with c-di-[ $^{32}$ P]GMP. **c**, **d**, HEK293T cells were transfected as in **a**, and the cell lysates were crosslinked with ultraviolet radiation to c-di-[ $^{32}$ P]GMP or [ $^{32}$ P]GTP in the presence of cold competing nucleotides in tenfold serial dilutions beginning at 1 mM (as indicated by the wedges), except

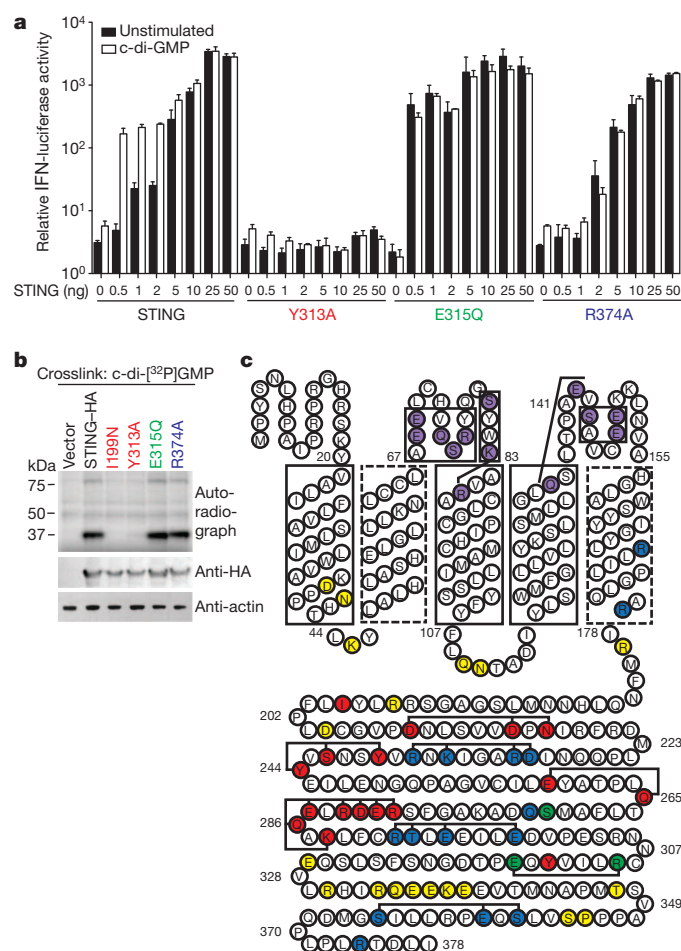
for guanosine (beginning at 0.1 mM), VV70mer (500  $\mu$ g ml<sup>-1</sup>), poly(dAT:dTA) (50  $\mu$ g ml<sup>-1</sup>) and poly(I)•poly(C) (denoted poly(I:C)) (50  $\mu$ g ml<sup>-1</sup>). The arrow indicates STING, and the asterisk indicates a nonspecific band. nt, nucleotides; ppGpp, guanosine-3',5'-bisphosphate. **e**, His<sub>6</sub>-STING (amino acids 138–378) (1  $\mu$ g) was separated by SDS-PAGE and stained with Coomassie blue. **f**, His<sub>6</sub>-STING 138–378 was analysed as in **c**. **g**, The binding of c-di-GMP to purified His<sub>6</sub>-STING (10  $\mu$ M) was measured by equilibrium dialysis.  $B_{\max}$ , maximum number of binding sites;  $h$ , hill coefficient;  $K_d$ , dissociation constant. **a–g**, Data are representative of three independent experiments.

dialysis to obtain an estimate of the affinity (dissociation constant,  $K_d$ ) of c-di-GMP binding to the STING CTD, which was  $\sim 5 \mu\text{M}$  (Fig. 2g). In its native membrane-bound form, or in complex with other host factors, STING may have a stronger affinity for c-di-GMP; nevertheless, a  $5 \mu\text{M}$  affinity is consistent with the dose response that has previously been observed in macrophages<sup>12</sup>. Consistent with the ability of STING to dimerize<sup>6</sup>, the binding data suggest a stoichiometry of one molecule of c-di-GMP per two molecules of STING.

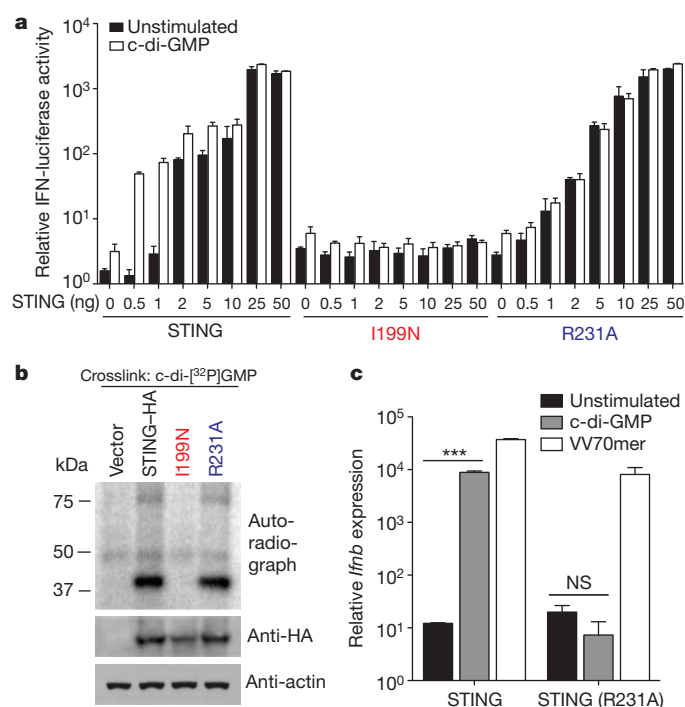
To identify the amino acids involved in c-di-GMP binding and/or IFN induction, we introduced point mutations into STING. Focusing on clusters of conserved and charged residues, we mutated 67 amino acids, either individually or in groups, and we classified these mutants into five categories (Fig. 3, Supplementary Table 1 and Supplementary Figs 3 and 4). Class I consists of mutants that abolish both binding and IFN induction (Fig. 3a–c, red, and Supplementary Table 1). Class II mutants bind to c-di-GMP but fail to induce IFN (Fig. 3c, purple). Class III comprises ‘hyperactive’ mutants, which spontaneously induce IFN at low levels of transfection (Fig. 3a–c, green, and Supplementary Table 1). Class IV mutants induce IFN when overexpressed but do not

respond to c-di-GMP (Fig. 3a–c, blue, and Supplementary Table 1). Class V consists of mutants that have no effect on c-di-GMP binding or IFN induction (Fig. 3c, yellow, and Supplementary Table 1). Although mutating STING can result in diverse phenotypes, a key finding is that all mutants that failed to bind to c-di-GMP also lost the ability to induce IFN in response to c-di-GMP. Consistent with our observation that the CTD is sufficient for binding to c-di-GMP (Fig. 2f), all mutations that affected c-di-GMP binding were located within the CTD.

DNA and cyclic dinucleotides induce indistinguishable transcriptional responses in macrophages<sup>10</sup>, and STING seems to be essential for both responses<sup>4,11</sup>. However, we found that STING expression is insufficient to restore the responsiveness of HEK293T cells to DNA, in contrast to cyclic dinucleotides (Fig. 1b and Supplementary Fig. 2a). Moreover, our competition assays indicate that DNA does not compete with cyclic-di-GMP for binding to STING under the conditions tested (Fig. 2d). Thus, although our data indicate that STING functions as a direct immunosensor of cyclic dinucleotides, it seems likely that additional host proteins are involved in IFN induction by DNA. Indeed, two candidate DNA sensors, DAI (also known as ZBP1) and IFI16, have been identified<sup>16,20</sup>, neither of which seems to be essential for the response to cyclic dinucleotides (ref. 10 and data not shown). To determine whether the responsiveness to cyclic dinucleotides and DNA are separable functions of STING, we sought to identify STING mutants that fail to respond to cyclic dinucleotides but still respond to DNA. We identified a STING mutant (R231A) that was unresponsive to c-di-GMP (Fig. 4a), although it induced IFN when overexpressed (Fig. 4a) and bound to c-di-GMP (Fig. 4b). Interestingly, STING R231A was able to restore the responsiveness of *gt* bone marrow macrophages to DNA but not to cyclic-di-GMP (Fig. 4c). Thus, cyclic



**Figure 3 | Mutational analysis of STING.** **a**, HEK293T cells were transfected with vectors encoding STING and STING mutants as indicated, together with an IFN-luciferase reporter, and luciferase activity was measured 6 h after stimulation. Data are presented as mean  $\pm$  s.d. ( $n = 3$ ). **b**, HEK293T cells were transfected as in **a**, except that the lysates were subjected to the crosslinking assay with c-di-[<sup>32</sup>P]GMP as in Fig. 2a. **c**, Organization of STING based on the membrane topology prediction programs SOSUI, TMHMM, HMMTOP and TMPred. Strongly predicted transmembrane domains are boxed; weakly predicted transmembrane domains have dashed boxes. Coloured residues indicate mutant classes (see text for description of mutants): class I (red), class II (purple), class III (green), class IV (blue) and class V (yellow). Bracketed mutations were made in combination. **a–c**, Data are representative of at least three independent experiments.



**Figure 4 | The IFN response to DNA and c-di-GMP can be uncoupled.** **a**, HEK293T cells were transfected as indicated, together with an IFN-luciferase reporter, and IFN reporter activity was measured 6 h after stimulation. **b**, HEK293T cells were transfected as in **a**, except that the lysates were subjected to the crosslinking assay with c-di-[<sup>32</sup>P]GMP as in Fig. 2a. **c**, Bone-marrow-derived macrophages from *Sting*-deficient (*gt*) mice were transduced with the indicated constructs. IFN induction in response to transfected cyclic-di-GMP or VV70mer was measured by quantitative RT-PCR and normalized to *Rps17*. \*\*\*,  $P < 0.001$ ; NS, not significant,  $P = 0.1205$ . **a, c**, Data are presented as mean  $\pm$  s.d. ( $n = 3$ ). **a–c**, Data are representative of at least three independent experiments.

dinucleotide sensing and DNA sensing can be uncoupled, suggesting that these two pathways are discrete but share STING as a common signalling molecule. It is unexpected that STING would function both as a direct immunosensor (of cyclic dinucleotides) and as a signalling adaptor (in the response to DNA). One possibility is that STING initially evolved as a cyclic dinucleotide sensor and was subsequently co-opted for DNA sensing.

We previously used mouse mutagenesis to identify STING as an essential molecule in the *in vivo* IFN response to cyclic dinucleotides<sup>11</sup>. The requirement for STING can now be rationalized by our proposal that STING functions as a direct sensor of cyclic dinucleotides. Interestingly, STING does not share homology with any known immunosensor and therefore seems to represent a novel category of microbial detector. Although a BLAST search of the mouse proteome for homologues of the *L. monocytogenes* diadenylate cyclase (Imo2120; also known as DacA) identifies STING as the top hit, the homology is limited to a short region of the STING CTD (amino acids 311–358). STING does not seem to share homology with PilZ-domain-containing proteins, which function as c-di-GMP receptors in bacteria<sup>8</sup>. Structural studies are required to better characterize the interaction of STING with cyclic dinucleotides and to determine whether STING resembles any known protein in mammals or bacteria.

Numerous studies have demonstrated that cyclic dinucleotides are potent immunostimulatory compounds that may be valuable as novel immunotherapeutics or adjuvants<sup>9,13</sup>. The therapeutic development of cyclic dinucleotides will be greatly facilitated by an improved understanding of the mechanism by which they are sensed. Furthermore, our finding that STING is a direct detector of cyclic dinucleotides provides insight into the fundamental mechanisms by which the innate immune system can detect bacterial infection.

## METHODS SUMMARY

**Transfections.** Transfections were carried out using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Digitonin permeabilization was used to introduce c-di-AMP into cells as described previously<sup>12</sup>.

**Recombinant STING.** DNA encoding the CTD of mouse STING (nucleotides 414–1,137) was cloned into the vector pET28a for recombinant protein expression in *Escherichia coli*.

**Ultraviolet radiation crosslinking.** c-di-[<sup>32</sup>P]GMP was enzymatically synthesized using recombinant WspR and was used in an ultraviolet radiation crosslinking assay as described previously<sup>21</sup>. Briefly, 50 µg HEK293T cell lysate at a final concentration of 2 µg µl<sup>-1</sup>, or 1 µg recombinant His<sub>6</sub>-tagged STING, was incubated with 2 µCi c-di-[<sup>32</sup>P]GMP in binding buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl and 1 mM MgCl<sub>2</sub>) for 15 min at 25 °C. The reactions were irradiated at 254 nm, and the proteins in the samples were then separated by SDS-PAGE.

**Statistical analysis.** Statistical differences were calculated with an unpaired two-tailed Student's *t*-test using Prism 5.0b software (GraphPad).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 29 April; accepted 3 August 2011.

Published online 25 September 2011.

1. Ishii, K. J. *et al.* A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nature Immunol.* **7**, 40–48 (2006).
2. Stetson, D. B. & Medzhitov, R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* **24**, 93–103 (2006).
3. Ishikawa, H. & Barber, G. N. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455**, 674–678 (2008).

4. Ishikawa, H., Ma, Z. & Barber, G. N. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461**, 788–792 (2009).
5. Jin, L. *et al.* MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals. *Mol. Cell. Biol.* **28**, 5014–5026 (2008).
6. Sun, W. *et al.* ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proc. Natl Acad. Sci. USA* **106**, 8653–8658 (2009).
7. Zhong, B. *et al.* The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity* **29**, 538–550 (2008).
8. Tamayo, R., Pratt, J. T. & Camilli, A. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annu. Rev. Microbiol.* **61**, 131–148 (2007).
9. Karaolis, D. K. *et al.* Bacterial c-di-GMP is an immunostimulatory molecule. *J. Immunol.* **178**, 2171–2181 (2007).
10. McWhirter, S. M. *et al.* A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP. *J. Exp. Med.* **206**, 1899–1911 (2009).
11. Sauer, J. D. *et al.* The *N*-ethyl-*N*-nitrosourea-induced *Goldenticket* mouse mutant reveals an essential function of *Sting* in the *in vivo* interferon response to *Listeria monocytogenes* and cyclic dinucleotides. *Infect. Immun.* **79**, 688–694 (2011).
12. Woodward, J. J., Iavarone, A. T. & Portnoy, D. A. c-di-AMP secreted by intracellular *Listeria monocytogenes* activates a host type I interferon response. *Science* **328**, 1703–1705 (2010).
13. Chen, W., Kuolee, R. & Yan, H. The potential of 3',5'-cyclic diguanylic acid (c-di-GMP) as an effective vaccine adjuvant. *Vaccine* **28**, 3080–3085 (2010).
14. Witte, G., Hartung, S., Buttner, K. & Hopfner, K. P. Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol. Cell* **30**, 167–178 (2008).
15. Janeway, C. A. Jr. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.* **54**, 1–13 (1989).
16. Unterholzner, L. *et al.* IFI16 is an innate immune sensor for intracellular DNA. *Nature Immunol.* **11**, 997–1004 (2010).
17. Ablasser, A. *et al.* RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nature Immunol.* **10**, 1065–1072 (2009).
18. Chiu, Y. H., Macmillan, J. B. & Chen, Z. J. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* **138**, 576–591 (2009).
19. Chodosh, L. A. in *Current Protocols in Molecular Biology* Ch. 12, 12.5.1–12.5.8 (Wiley, 2001).
20. Takaoka, A. *et al.* DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* **448**, 501–505 (2007).
21. Merighi, M., Lee, V. T., Hyodo, M., Hayakawa, Y. & Lory, S. The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **65**, 876–895 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank H. Carlson, K. Collins, S. McWhirter, D. Raulet, K. Sjölander and members of the Vance, Barton and Portnoy laboratories at the University of California, Berkeley, for advice and discussions. We thank J. Woodward and D. Portnoy for their gift of purified c-di-AMP. Work in R.E.V.'s laboratory is supported by investigator awards from the Burroughs Wellcome Fund and the Cancer Research Institute and by National Institutes of Health (NIH) grants AI075039, AI080749 and AI063302. D.L.B. is supported by an NIH National Research Service Award fellowship F32 (AI091100).

**Author Contributions** D.L.B. performed luciferase assays and quantitative RT-PCR, generated c-di-[<sup>32</sup>P]GMP, purified recombinant STING, performed c-di-[<sup>32</sup>P]GMP binding assays and transduced *gt* bone marrow macrophages. K.M.M. generated truncation mutations and performed luciferase assays. K.S.-T. generated point mutants and performed luciferase assays. D.L.B., K.M.M. and R.E.V. participated in study design and data analysis. D.L.B. and R.E.V. wrote the paper. B.E. contributed to protein purification methods. J.S.I. contributed to the design of the equilibrium dialysis experiments and the analysis of binding data. M.H. and Y.H. synthesized c-di-GMP.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to R.E.V. ([rvance@berkeley.edu](mailto:rvance@berkeley.edu)).

## METHODS

**Synthesis of c-di-[ $^{32}$ P]GMP.** The synthesis of c-di-[ $^{32}$ P]GMP was carried out as described previously<sup>21</sup>. Briefly, recombinant His<sub>6</sub>-tagged WspR was incubated with [ $\alpha$ - $^{32}$ P]GTP (3,000 Ci mmol<sup>-1</sup>, 10  $\mu$ Ci  $\mu$ l<sup>-1</sup>, Amersham Biosciences) for 2 h at 25 °C, followed by heat inactivation of WspR at 95 °C for 5 min. Residual  $^{32}$ P was removed by incubation with calf intestinal phosphatase (CIP) (New England Biolabs) for 10 min at 37 °C. CIP was heat inactivated at 95 °C for 5 min, followed by centrifugation at 16,000g for 5 min. The [ $^{32}$ P]GTP used as a negative control was prepared identically except that His<sub>6</sub>-tagged WspR and CIP were omitted from the preparation. Radiolabelled nucleotides were quantified by separation by thin-layer chromatography on cellulose-PEI plates (Macherey-Nagel) using 1.5 M KH<sub>2</sub>PO<sub>4</sub>, pH 3.65 (Supplementary Fig. 5).

**Cell lines and animals.** C57BL/6 *Myd88*<sup>-/-</sup>*Trif*<sup>-/-</sup> mice were obtained from G. Barton, and immortalized macrophages were generated as described previously<sup>22</sup>. Immortalized bone marrow macrophages were maintained in RPMI 1640 (Invitrogen) supplemented with 10% FBS, penicillin-streptomycin and glutamine. HEK293T cells were maintained in DMEM supplemented with 10% FBS, penicillin-streptomycin and glutamine. Animal use was approved by the Animal Care and Use Committee at the University of California, Berkeley.

**Plasmids.** A construct encoding RocR (NP\_252636) from *P. aeruginosa*<sup>23</sup> was a gift from S. Lory. *rocR* was cloned into the MSCV2.2 retroviral expression construct upstream of an internal ribosome entry site (IRES)-green fluorescent protein (GFP). MSCV-*rocR* was transduced into immortalized macrophages from *Myd88*<sup>-/-</sup>*Trif*<sup>-/-</sup> mice, and cells were sorted for GFP expression. Mouse *Sting* and the *gt* (I199N) mutant allele of *Sting* were cloned into the vector pcDNA3 with a C-terminal HA tag as described previously<sup>11</sup>. DNA encoding the CTD of mouse *Sting* (nucleotides 414–1,137) was cloned into pET28a for recombinant protein expression in *Escherichia coli*.

**Site-directed mutagenesis.** Mutations in *Sting* were generated using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's guidelines.

**Reagents.** c-di-GMP was synthesized as described previously<sup>24</sup>. Purified c-di-AMP was a gift from J. Woodward and D. Portnoy. Poly(dAT:dTA), GTP, ATP, GMP and guanosine were obtained from Sigma-Aldrich. Poly(I)\*poly(C) (denoted poly(I:C)) was purchased from Invivogen. Guanosine-3',5'-bisdiphosphate (ppGpp) was obtained from TriLink Biotechnologies. Sendai virus was purchased from Charles River Laboratories. Theiler's murine encephalomyelitis virus (TMEV) strain GDVII was provided by M. Brahic and E. Freundt. DNA oligonucleotides corresponding to the VV70mer and ISD were purchased from Elim Biopharmaceuticals and were annealed as described previously<sup>2,16</sup>.

**Cell stimulation.** All transfections (excluding for c-di-AMP) were carried out using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. The VV70mer was transfected at a final concentration of 0.5  $\mu$ g ml<sup>-1</sup>. Poly(dAT:dTA), poly(I:C) and c-di-GMP were transfected at a final concentration of 4  $\mu$ g ml<sup>-1</sup>. c-di-AMP was used at a final concentration of 5.4 mM, and stimulation was performed using digitonin permeabilization as described previously<sup>12</sup>.

**Luciferase assay.** HEK293T cells were plated in TC-treated 96-well plates at  $0.5 \times 10^6$  cells ml<sup>-1</sup>. The next day, the cells were transfected as indicated, together with IFN- $\beta$ -firefly luciferase and TK-*Renilla* luciferase reporter constructs. Following stimulation for 6 h with the indicated ligands, the cells were lysed in passive lysis buffer (Promega) for 5 min at 25 °C. The cell lysates were incubated with firefly luciferase substrate (Biosynth) and the *Renilla* luciferase substrate coelenterazine (Biotium), and luminescence was measured on a SpectraMax L microplate reader (Molecular Devices). The relative *Irfn* expression was calculated as firefly luminescence relative to *Renilla* luminescence.

**Quantitative PCR.** The analysis of *Irfn* expression by bone marrow macrophages was conducted as described previously<sup>10</sup>.

**Preparation of HEK293T cell lysates and immunoprecipitations.** HEK293T cells were plated at a density of  $1 \times 10^6$  cells well<sup>-1</sup> in a 6-well plate. The following day, the cells were transfected with pcDNA3 or pcDNA3 expressing HA-tagged

wild-type or mutant STING using Lipofectamine 2000 (Invitrogen). The day after that, the cells were rinsed once with PBS and transferred to Eppendorf tubes in PBS containing 1 mM EDTA. The cells were pelleted briefly by centrifugation at 1,000g at 4 °C. The cell pellet was lysed in an equal volume of digitonin lysis buffer (0.5% digitonin, 20 mM Tris-HCl, pH 7.4, and 150 mM NaCl) containing protease inhibitors (Roche) for 10 min on ice. The cell lysates were centrifuged at 10,000g for 10 min at 4 °C. The protein concentration in the resultant supernatant was measured using the Bradford reagent (Bio-Rad). The cell lysates were subjected to a c-di-GMP binding (crosslinking) assay (see below). The lysates were separated by SDS-PAGE, and the separated proteins were transferred to a nitrocellulose membrane, which was then probed with rat anti-HA antibodies (Roche), to confirm STING-HA expression, and mouse anti- $\beta$ -actin antibodies (Santa Cruz Biotechnology). To immunoprecipitate HA-tagged STING, the cell lysates were prepared similarly in digitonin lysis buffer and incubated with anti-HA-antibody-conjugated agarose beads (Sigma) for 2 h at 4 °C. Washed beads were subjected to a c-di-GMP binding assay or separated by SDS-PAGE and stained with colloidal blue protein stain (Thermo Scientific).

**c-di-GMP binding assays.** The c-di-GMP binding assay (also called the crosslinking assay) was based on a method described previously<sup>21</sup>. Briefly, 50  $\mu$ g HEK293T cell lysate at a final concentration of 2  $\mu$ g  $\mu$ l<sup>-1</sup>, or 1  $\mu$ g recombinant His<sub>6</sub>-tagged STING, was incubated with 2  $\mu$ Ci radiolabelled nucleotide in binding buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl and 1 mM MgCl<sub>2</sub>) for 15 min at 25 °C. The reactions were irradiated at 254 nm for 20 min on ice at a 3-cm distance from a UVG-54 mineral light lamp (UVP). Immediately after crosslinking, the reactions were terminated by the addition of SDS sample buffer (40% glycerol, 8% SDS, 2% 2-mercaptoethanol, 40 mM EDTA, 0.05% bromophenol blue and 250 mM Tris-HCl, pH 6.8), boiled for 5 min and then separated by SDS-PAGE. The gels were dried, exposed to a phosphor screen and visualized using a Typhoon Trio imager (GE Healthcare).

**Protein purification.** The construct expressing a constitutively active form of WspR (pQE-WspR\*)<sup>23,25</sup> was a gift from S. Lory. The purification of His<sub>6</sub>-tagged WspR was carried out as described previously, using Ni-NTA affinity chromatography (Qiagen)<sup>26</sup>. DNA encoding the CTD of mouse STING (nucleotides 414–1,137) was cloned into the vector pET28a and purified by Ni-NTA affinity chromatography (Qiagen) according to the manufacturer's instructions.

**Equilibrium dialysis.** The binding affinity of radioactive c-di-GMP was measured by equilibrium dialysis, using a 96-well equilibrium dialyser (Harvard Apparatus) with a 5,000 molecular weight cut-off membrane. One chamber contained 150  $\mu$ l 10  $\mu$ M purified His<sub>6</sub>-tagged STING(138–378) in assay buffer (25 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM MgCl<sub>2</sub> and 10% glycerol), while the other was filled with 150  $\mu$ l c-di-[ $^{32}$ P]GMP at a range of concentrations (40–160  $\mu$ M). Equilibrium was reached after 48 h at 25 °C, and three samples were drawn from each chamber and mixed with 2 ml Econo-Safe scintillation fluid. Samples were measured in an LS 6000 IC scintillation counter (Beckman). Data analysis was performed using Prism 5.0b software (GraphPad). The dissociation constant ( $K_d$ ), the maximum number of binding sites ( $B_{max}$ ) and the Hill coefficient ( $h$ ) were generated using nonlinear regression, allowing one-site specific binding with a Hill slope.

22. Broz, P., von Moltke, J., Jones, J. W., Vance, R. E. & Monack, D. M. Differential requirement for caspase-1 autoproteolysis in pathogen-induced cell death and cytokine processing. *Cell Host Microbe* **8**, 471–483 (2010).
23. Kulesekara, H. et al. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc. Natl Acad. Sci. USA* **103**, 2839–2844 (2006).
24. Hyodo, M. & Hayakawa, Y. An improved method for synthesizing cyclic bis-(3'-5')diguanylic acid (c-di-GMP). *Bull. Chem. Soc. Jpn* **77**, 2089–2093 (2004).
25. Simm, R., Remminghorst, U., Ahmad, I., Zakikhany, K. & Romling, U. A role for the EAL-like protein STM1344 in regulation of CsgD expression and motility in *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **191**, 3928–3937 (2009).
26. De, N. et al. Phosphorylation-independent regulation of the diguanylate cyclase WspR. *PLoS Biol.* **6**, e67 (2008).

# Temporal dynamics and genetic control of transcription in the human prefrontal cortex

Carlo Colantuoni<sup>1,2,3,4\*</sup>, Barbara K. Lipska<sup>1\*</sup>, Tianzhang Ye<sup>1</sup>, Thomas M. Hyde<sup>1,4</sup>, Ran Tao<sup>1</sup>, Jeffrey T. Leek<sup>2</sup>, Elizabeth A. Colantuoni<sup>2</sup>, Abdel G. Elkahouloun<sup>5</sup>, Mary M. Herman<sup>1</sup>, Daniel R. Weinberger<sup>1,4</sup> & Joel E. Kleinman<sup>1</sup>

Previous investigations have combined transcriptional and genetic analyses in human cell lines<sup>1–3</sup>, but few have applied these techniques to human neural tissue<sup>4–8</sup>. To gain a global molecular perspective on the role of the human genome in cortical development, function and ageing, we explore the temporal dynamics and genetic control of transcription in human prefrontal cortex in an extensive series of post-mortem brains from fetal development through ageing. We discover a wave of gene expression changes occurring during fetal development which are reversed in early postnatal life. One half-century later in life, this pattern of reversals is mirrored in ageing and in neurodegeneration. Although we identify thousands of robust associations of individual genetic polymorphisms with gene expression, we also demonstrate that there is no association between the total extent of genetic differences between subjects and the global similarity of their transcriptional profiles. Hence, the human genome produces a consistent molecular architecture in the prefrontal cortex, despite millions of genetic differences across individuals and races. To enable further discovery, this entire data set is freely available (from Gene Expression Omnibus: accession GSE30272; and dbGaP: accession phs000417.v1.p1) and can also be interrogated via a biologist-friendly stand-alone application (<http://www.libd.org/braincloud>).

The temporal dynamics of genome expression throughout the body and its genetic and epigenetic control are central to a synthetic understanding of how a relatively small number of DNA molecules can give rise to an entire human. Similarly, temporal expression patterns in neural tissue and their regulation across the lifespan will elucidate molecular mechanisms involved in the formation, mature function and degeneration of the human brain.

Previous studies have combined transcriptome and genetic analyses to investigate the genetic control of gene expression in human cell lines<sup>1–3</sup>. Few studies have applied these genomic techniques to human neural tissue<sup>4–7</sup> or human brain disease<sup>8</sup>. Others have focused on the transcriptome in human fetal brain tissue<sup>9</sup>, temporal patterns of gene expression in postnatal life<sup>10</sup>, and gene co-expression patterns in the brain<sup>11,12</sup>. Here we describe the combination of genome-wide DNA and RNA analyses in a large collection of meticulously curated human brain specimens to produce a comprehensive view of how the expression of the human genome in the prefrontal cortex (PFC) progresses from fetal development through ageing and how sequence variation in the genome impacts on these expression patterns.

The post-mortem brain tissue collection ( $n = 269$  subjects without neuropathological or neuropsychiatric diagnosis) spans the majority of the human lifespan (Fig. 1a, b). From each subject in the brain collection, RNA from PFC grey matter was analysed using spotted oligonucleotide microarrays yielding data from 30,176 gene expression probes. DNA from cerebellar tissue was studied with Illumina BeadChips producing 625,439 single nucleotide polymorphism (SNP) genotypes for each subject.

The absolute rate of expression change within each life stage was quantified for all genes using linear models (Fig. 1b, box plot). The rate of expression change during fetal development is much faster than at any other stage in human life. Changes during infancy are much slower, yet still more rapid than at any later time in life. After the first half year of postnatal life, rates of expression change slow markedly, and continue to slow during the childhood and teenage years, subsequently maintaining a low rate of change through the 20s, 30s and 40s. After this period, rates of expression change begin to rise again through several decades, and in the aged human brain, change reaches and then exceeds rates observed during teenage years.

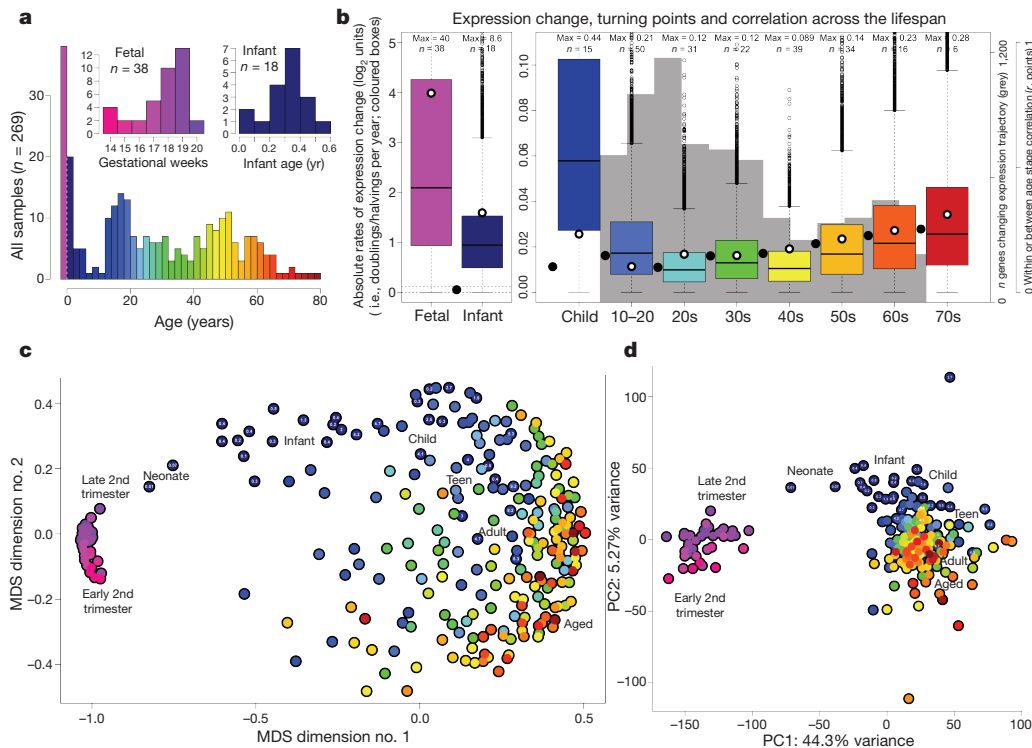
The distribution of expression trajectory turning points was investigated across postnatal life (Fig. 1b, grey histogram). Rates of expression change decrease from childhood through the teenage years (blue boxes) as many genes redirect expression trajectories (peak in grey histogram near 20 years). In contrast, in ageing, expression change accelerates (yellow–orange boxes) as more genes enter turning points (the minor peak in the grey histogram near 60 years).

The correlation of expression measures across subjects was explored within each age stage and between adjacent stages (Fig. 1b, points). Transcription in PFC appears most similar across individuals at the beginning of life and then again to a lesser extent nearer its end, demonstrating the most diversity during the years of mature brain functioning, when age-dependent rates of expression change are lowest (this observation is also clear in Fig. 1c). The separation of mean within- and between-age stage correlations observed early in life indicates the occurrence of fundamentally distinct transcriptional programs within fetal, infant and childhood development, followed by a smoother more continuous progression of change throughout the rest of the lifetime.

To obtain a global perspective on transcription in PFC across the human lifetime, expression profile correlations were combined with multidimensional scaling (MDS) to reduce the complexity of the expression data and produce an intuitive visualization of global patterns (Fig. 1c). The spatial progression of the colour scale in this plot is a reflection of age-dependent change in human PFC transcription. Even within the brief 6-week range of fetal development examined, there is clearly observable systematic expression change with time (along the vertical axis). Following fetal development, the path of global transcriptional change alters markedly, progressing steadily away from the fetal state through the neonatal, infant and childhood ages, each of which has a relatively distinct identity compared with other periods (across the horizontal axis). A second redirection of global transcriptional change occurs at the end of the teenage years (also observed from a different perspective in Fig. 1b, grey histogram), followed by a more linear progression through adulthood and into ageing. This global view was also used to inspect the effects of covariates (Supplementary Fig. 1).

<sup>1</sup>Section on Neuropathology, Clinical Brain Disorders Branch, Genes, Cognition and Psychosis Program, IRP, NIMH, NIH, Bethesda, Maryland 20892, USA. <sup>2</sup>Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA. <sup>3</sup>Illumina Biotechnology, Inc., Baltimore, Maryland 21211, USA. <sup>4</sup>The Lieber Institute for Brain Development, Johns Hopkins University Medical Center, Baltimore, Maryland 21205, USA. <sup>5</sup>Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

\*These authors contributed equally to this work.



**Figure 1 | A global view of the PFC transcriptome.** **a**, Histogram of subject ages in the brain collection. Birth is indicated by the dashed white line. This colour scale serves as the legend for all panels in Fig. 1 and Fig. 2b. **b**, Box plot of absolute rates of expression change within each stage of life. Because the rates of expression change are so high early in life, the y-axis scale is different for fetal and infant stages than for all other stages. The two horizontal dotted lines in the left panel show the entire extent of the y axis in the right panel. Only microarray probes showing systematic variation with age ( $R^2 > 0.5$ ) were included (n = 8,704 probes). Age ranges: fetal, 14–20 gestational weeks; infant, 0–6 months; child, 1–10 years; decades as labelled. Open points depict the mean expression correlation across subjects within each age stage (Pearson's  $r$  calculated across all expression measures; y axis scale at far right). Filled points

In another global view of prefrontal transcription, the age effect within the fetal samples is effectively illustrated using principal components analysis (PCA, Fig. 1d). The first principal component (PC1) separates the fetal from postnatal samples, whereas the second (PC2) appears to align with age effects within both the fetal and postnatal samples. The directions of the fetal and postnatal age effects along PC2 appear to be in opposition. Additionally, fetal expression changes are negatively correlated with those in other stages of early life: infancy  $r = -0.45$ ,  $P = 1.3 \times 10^{-90}$ ; childhood  $r = -0.48$ ,  $P = 1.5 \times 10^{-47}$ ; and teenage years  $r = -0.18$ ,  $P = 2.3 \times 10^{-8}$  (including only probes with slopes at  $P < 0.05$  in both stages, Supplementary Table 1). This might indicate that select fetal expression changes are reversed at different times across the lifespan, beginning immediately after birth.

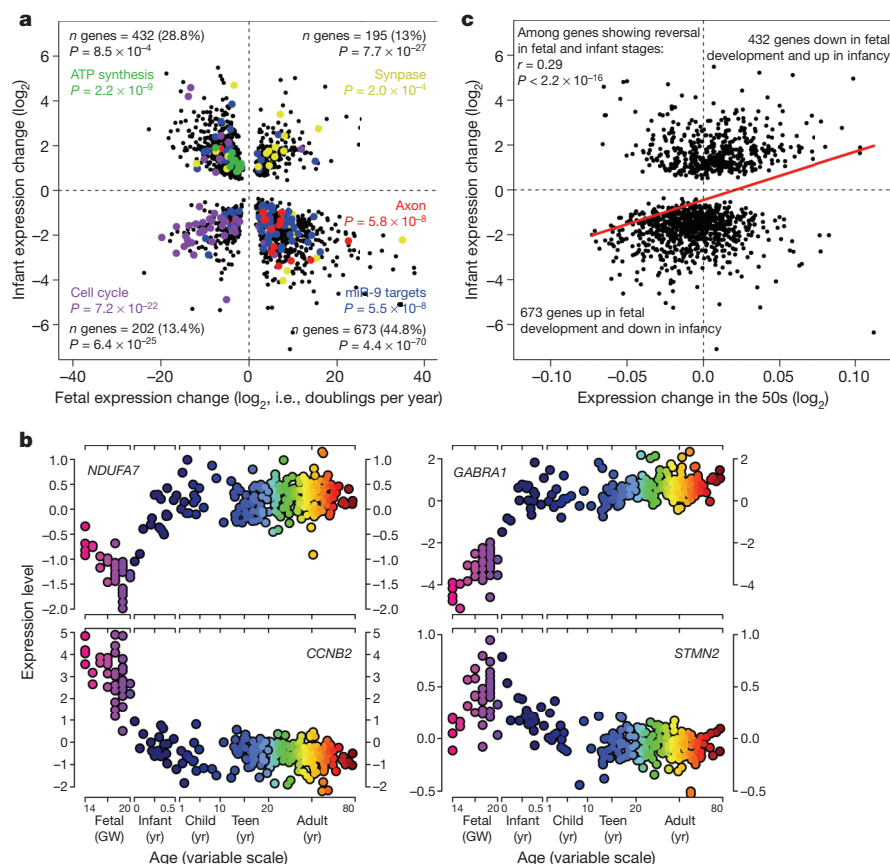
To investigate further this observation of reversing trajectories, genes showing significant expression change across age in both fetal and infant development were compared directly (Fig. 2a, b). Approximately three-quarters of genes showing significant change in both stages reverse their direction of expression change between fetal and early postnatal life, with most changing from an increase *in utero* to a decrease in the months after birth.

To gain functional insight into these changing expression patterns, the genes within each of the quadrants in Fig. 2a were interrogated for the over-representation of functional gene groups. Detailed functional group lists for each of the quadrants are contained in Supplementary Table 3. This examination of gene expression trajectories in early life may give a global genomic perspective on mechanisms in neural

development that have been well studied at the individual cell and gene level: genes involved in cell division are over-represented among genes for which expression decreases during both fetal development and infancy. Inversely, genes related to the synapse are over-represented among genes showing expression increases during both stages. This pair of findings is a genomic reflection of the well-characterized decrease in cell proliferation with opponent increase in neuronal differentiation through both late fetal and early infant development.

In contrast to synaptic components, genes with axonal function are highly enriched among genes showing increasing expression during fetal development followed by decreases after birth. This coordinated reversal of expression trajectories among axonal genes while many synaptic genes continue to increase in infancy is probably a genomic view of the process of pruning exuberant axons while synapse development and maturation at appropriate target sites advance<sup>13</sup>. Specific gene expression changes in synaptic and axonal genes during fetal and infant life are listed in Supplementary Table 4.

Genes in ATP synthesis also show a reversal of expression patterns, but in this case, decreasing during fetal development and rising after birth. In fetal development, energy metabolism seems to be slowing along with the decrease in cellular proliferation, consistent with cell division as the primary energy consuming process during fetal development. However, after birth, proliferation in the PFC continues to slow while expression of energy metabolism genes increases markedly. Other functional gene groups with increasing expression during these first postnatal months include genes involved in  $\text{Ca}^{2+}$  binding,  $\text{Ca}^{2+}$



**Figure 2 | Reversal of fetal expression changes in infancy and ageing.**

**a**, Scatter plot of fetal and infant gene expression change with age. Each gene is represented as a single point. Only genes with slopes at  $P < 0.05$  in both stages were included in this analysis ( $n = 1,502$  genes measured by 1,819 probes; Supplementary Table 2). The number of genes in each quadrant is indicated in black. The  $P$ -values listed were derived from Pearson's  $\chi^2$  tests comparing the proportion of genes in each quadrant compared to an expected proportion of 0.25. Key functional gene groups are highlighted and listed in the quadrants where they are over-represented. **b**, Depiction of individual genes' expression across the human lifespan, illustrating the four patterns of expression across fetal and infant development shown in **a**. The gene depicted in each panel is an

example drawn from gene groups highlighted in the corresponding quadrants in **a**. Expression levels (normalized  $\log_2$ (sample/reference)) are on the y axis, with age on the x axis. The linear age scale is variable within each of the individual stages of life as labelled. Subjects are coloured by age as in Fig. 1a. GW, gestational week. Clockwise from top left: NADH dehydrogenase (ubiquinone) 1 $\alpha$  subcomplex 7, 14.5 kDa (*NDUFA7*; ATP synthesis);  $\gamma$ -aminobutyric acid A receptor,  $\alpha 1$  (*GABRA1*; synapse); stathmin-like 2 (*STMN2*; axon); and cyclin B2 (*CCNB2*; cell cycle). **c**, Visualization of ageing (50s) and infant expression trajectories among genes showing reversal between fetal development and infancy (same genes depicted in **a**, lower right and upper left panels,  $n = 1,105$ ). The red line depicts a linear fit to the data.

transport, gated ion channels, voltage-gated  $K^+$  channels and active ion transport (Supplementary Table 3), indicating that neuronal maturation and activity now drive energy production.

This functional analysis of expression trajectories also reveals potentially novel mechanisms in early cortical development: in the heavily populated quadrant showing increasing expression in the fetus and decreasing expression in infancy, 22 of the top 49 over-represented gene groups are microRNA (miRNA) target gene groups (Supplementary Table 3,  $P = 6.5 \times 10^{-5}$  and below). Together, these miRNA target groups account for 266 of the 673 genes in this quadrant (40%). miR-9 targets are the most highly enriched of these miRNA target gene groups. miR-9 is brain-specific<sup>14</sup> and is used reiteratively in diverse processes in neural development, including patterning, neurogenesis and differentiation<sup>15,16</sup>, as well as cell migration<sup>17</sup>.

The reversal of fetal expression trajectories is also seen much later in life. Fetal expression trajectories show a strong negative correlation with changes observed in the sixth decade of life (50s) ( $r = -0.46$ ,  $P = 2.4 \times 10^{-21}$ ; Supplementary Table 1). This finding is consistent with the age-dependent repression of neuronal genes observed previously<sup>18</sup>. Whereas fetal expression trajectories show negative correlation with both infant and 50s trajectories, expression trajectories in infancy do not correlate with those observed in the 50s (Supplementary Table 1). However, within the set of genes showing trajectory

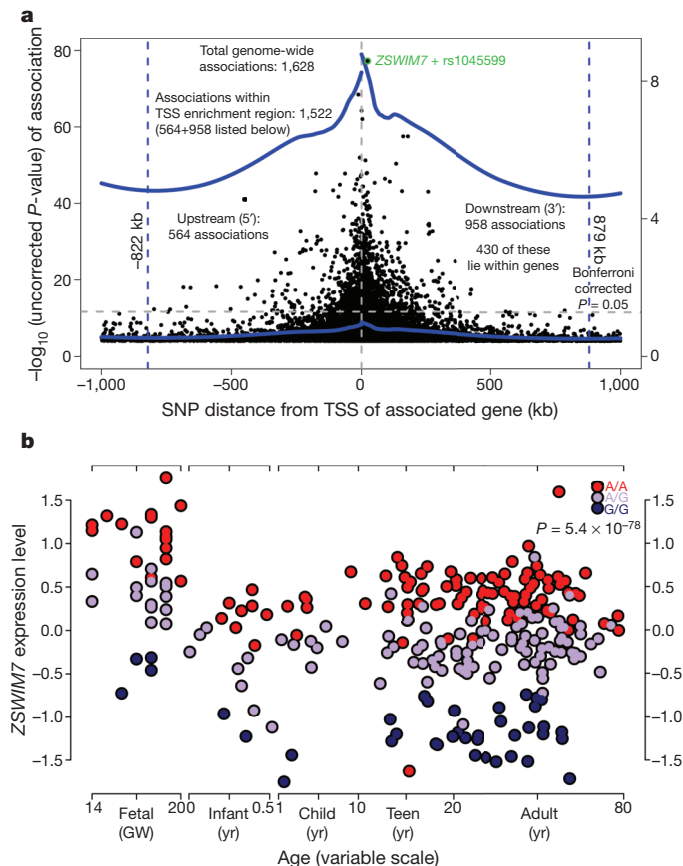
reversal between fetal and infant ages, expression change in infancy and in the 50s share a striking amount of similarity (Fig. 2c). Therefore, although infant expression changes do not globally resemble those happening later in life, the specific reversal of fetal expression trajectories seen in infancy is mirrored within changes in ageing.

These fetal reversals in ageing can also be demonstrated by comparing our observations in fetal development with recent findings in ageing. Genes with significant increases during fetal development are enriched for genes shown to decrease in the ageing cortex<sup>19</sup>, whereas genes decreasing during fetal development are enriched for genes known to increase in ageing ( $P = 1.0 \times 10^{-6}$  and  $P = 4.6 \times 10^{-11}$ , respectively; see Supplementary Table 5). Similar reversals are also seen in genes reported to change in Alzheimer's disease<sup>20</sup>: fetal increases are enriched for genes downregulated in Alzheimer's disease and fetal decreases are enriched for genes upregulated in Alzheimer's disease ( $P = 2.2 \times 10^{-21}$  and  $P = 7.1 \times 10^{-7}$ , respectively; see Supplementary Table 5). Hence, in the PFC, the reversal of specific expression patterns from *in utero* development occurs in infancy and then again much later in normal ageing and in the neuropathological processes of Alzheimer's disease.

To explore the genetic control of prefrontal expression patterns, DNA from the sample collection was interrogated with high-density SNP microarrays to catalogue common genomic polymorphisms.

All possible associations of SNP genotypes with gene expression levels were examined (expression quantitative trait loci, or eQTL):  $n = 30,176$  expression probes  $\times 625,439$  SNP genotypes  $= 1.89 \times 10^{10}$  ( $\sim 19$  billion) possible associations. Consistent with previous observations, we see that individual SNPs can profoundly affect the expression level of individual genes. When considering data across all subjects, 1,628 individual associations surpass genome-wide Bonferroni correction. Association analysis was also conducted within the African American and Caucasian samples separately (significant associations for all analyses are in Supplementary Table 6).

The strength and location of associations relative to transcriptional start sites (TSS) are explored in Fig. 3a. Consistent with past eQTL studies across many organisms, we find that effects proximal to TSSs are of greater average strength than associations across greater



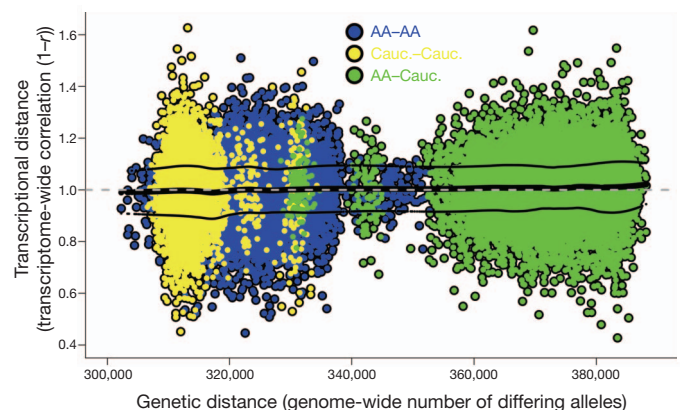
**Figure 3 | Genetic control of PFC gene expression.** **a**, Position and strength of associations of SNP genotype with gene expression levels (distance of the SNP from the associated TSS is plotted on the x axis). Only gene-SNP pairs lying on the same chromosome are displayed here. Negative log<sub>10</sub>(P value) is on the y axis. Only P values  $< 0.0001$  are included in this analysis. The genome-wide Bonferroni-corrected  $P = 0.05$  is shown as a horizontal dashed grey line:  $P = 0.05/1.89 \times 10^{10}$  associations  $= 2.6 \times 10^{-12}$  (11.6 on this y-axis scale). The numbers of associations listed in the plot refer to those passing this genome-wide alpha level. Both solid blue curves depict a local nonlinear regression fit (loess, span = 0.5) of association strength across distance from the TSS. The blue fit lying close to the y axis is a fit plotted in the same scale as the plotted P values (y axis at left). The second blue fit is the same fit, plotted on an expanded y axis (y axis at right). The minimum value of this fit both upstream and downstream from the TSS (that is, approximately where the local estimate of enrichment for greater association reaches zero) is marked with a vertical blue dashed line. **b**, The most significant observed association of a single SNP with the expression of one gene across the lifespan (highlighted in **a** in green): the rs1045599 SNP lies within the ZSWIM7 gene on chromosome 17 (ZSWIM7, zinc finger, SWIM-type containing 7). Age scales are defined as in Fig. 2b. ZSWIM7 expression level (normalized log<sub>2</sub>(sample/reference)) is on the y axis and is coloured by rs1045599 genotype.

distances or across chromosomes (Fig. 3a). There are considerably more strong associations downstream (3') from the TSS than upstream. This is consistent with previous observations<sup>8</sup>, and demonstrates that downstream polymorphisms (often within gene sequences) that impact on expression are stronger and/or more numerous than alterations at equal distances upstream (potentially in promoter or enhancer sequences). Additionally, expression-associated SNPs are biased towards positions within genes (fold enrichment = 1.61,  $P = 2.9 \times 10^{-76}$ ). Within this gene bias, both exonic and intronic locations are over-represented, but to vastly different degrees (fold enrichment = 4.3 and 1.4,  $P = 5.0 \times 10^{-94}$  and  $1.2 \times 10^{-32}$ , respectively).

The single strongest association observed was between the expression of the ZSWIM7 gene and SNP rs1045599, located within this same gene (Fig. 3a, b). This association of genotype with expression level is observed across all ages and races studied. Similar to this analysis, the freely available interactive stand-alone application that we have developed enables the visualization of expression data across the lifespan and the exploration of genetic associations for individual gene queries (<http://www.libd/braincloud>). We invite the research community to explore this resource with their own interests.

To explore the relationship between the genome as a whole and the PFC transcriptome as a whole, we compared genetic distance and transcriptional distance in all possible pairwise subject comparisons (Fig. 4). Although individual SNPs clearly have an impact on the expression of individual genes (Fig. 3 and Supplementary Table 6) globally, there is no association of genetic distance between individual humans with the similarity of their prefrontal transcriptional profiles (Fig. 4,  $R^2 = 0.002$ ).

This dramatic lack of association between genetic distance and transcriptome distance across our sample is a surprising result that requires further interrogation. It is possible that no association is found in Fig. 4 because most of the genetic polymorphisms measured do not impact on gene expression. Therefore, we repeated this search for association by investigating global transcriptional distance across a focused subset of the genetic data: only SNPs involved in genome-wide significant SNP-expression associations were considered. This analysis also revealed no association between focused genetic distance



**Figure 4 | The genome produces a consistent molecular architecture in PFC.** Global comparison of genetic and transcriptional differences between subjects. Each point represents a comparison of two subjects in the collection. Genetic distance between subjects is depicted on the x axis as the number of differing alleles over the portion of the genome interrogated. Transcriptional distance is shown on the y axis as 1 minus the correlation across all gene expression values from the subjects (as used in Fig. 1c). Each subject comparison is coloured to indicate the races (AA, African American; Cauc., Caucasian) of the two individuals involved in the comparison. The thick black curve is an estimate of the local mean (loess, span = 0.25) of transcriptional distance as it varies across genetic distance. The thin black curves depict fits to the residuals around this mean. Only African American and Caucasian sample comparisons are visualized here (>96% of the collection).

and global transcriptional distance (Supplementary Fig. 2). In addition, these same analyses performed within individual races showed no association between global transcriptional distance and genetic distance when either global or focused genetic distance was used.

We conclude that despite the many genetic polymorphisms that individually can affect the expression of single genes, the human genome produces a consistent molecular architecture in the human prefrontal cortex across the lifespan. This is true across (the human) race. It is possible that individual genetic traits and complex combinations of traits that disrupt this architecture are selected against in the general population and would not appear in studies of normal human brain development. The clear observation of associations of individual genetic polymorphisms with gene expression (Fig. 3) in the absence of a relationship between global genetic and transcriptome profiles (Fig. 4) demonstrates our ability to analyse microscale genetic effects while macroscale interactions remain elusive. It is perhaps useful to consider each individual complete genome as a grand combination of variants which is acted upon (in evolution and in environment) and which acts (in development, biological function and disease) as a whole, rather than individual genetic traits in isolation. Characterization of the higher-order interactions within this whole is a great challenge facing biologists today.

By creating this freely available public resource, we hope that the research community can further explore this data set. This full data set is downloadable at <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30272> (expression data) and [http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study\\_id=phs000417.v1.p1](http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000417.v1.p1) (SNP data). In addition, we have constructed a biologist-friendly stand-alone application designed to allow the research community to interrogate this resource one gene at a time (<http://www.libd.org/braincloud>).

## METHODS SUMMARY

**Brain tissue collection.** Post-mortem human brains from the NIMH Brain Tissue Collection in the Clinical Brain Disorders Branch (NIMH, CBDB) were obtained at autopsy. Additional brain tissue samples were provided by the NICHD Brain and Tissue Bank for Developmental Disorders (<http://www.BTBank.org>). Clinical characterization, diagnoses and macro- and microscopic neuropathological examinations, toxicological analysis, RNA extraction and quality control measures were performed using a standardized paradigm. Subjects with evidence of neuropathology, drug use, alcohol abuse, or psychiatric illness were excluded. Subject demographics and sample details are contained in Supplementary Table 7.

**RNA resources.** Post-mortem PFC grey matter tissue homogenates were obtained from all subjects. Total RNA was extracted, amplified and fluorescently labelled. Reference RNA was pooled from all samples and treated identically to sample RNAs. Labelled RNAs were hybridized to two-colour custom-spotted arrays from the NHGRI microarray core facility. After normalization<sup>21</sup>, log<sub>2</sub> intensity ratios were further adjusted to reduce the impact of known and unknown sources of systematic noise on gene expression measures using surrogate variable analysis<sup>22</sup> (SVA). Validation of microarray expression patterns was performed by Taqman qPCR (Supplementary Table 8). In this study of RNA derived from tissue homogenates, differential gene expression within a population of cells stable in cell type is indistinguishable from a change in the abundance of cell types that express different genes. There is no doubt that both phenomena contribute to signals measured here in the prefrontal cortex.

**DNA resources.** DNA for genotyping was obtained from the cerebella of samples in the collection and applied to Illumina BeadArrays. Genotypes were called using BeadExpress software.

**Functional gene groupings.** To generate functional gene groups for the analysis described in Fig. 2a and the text, microarray probes were annotated with data from numerous public, online sources. Enrichment of functional gene groups within various gene lists as described in the text was assessed by a standard hypergeometric test.

**SNP-expression associations.** SNP-expression associations were carried out using linear models that included surrogate variables, age, life stage (as defined in Fig. 1b), an interaction of age and life stage, sex, race and the SNP under investigation. SNP was included as a continuous variable.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 29 November 2010; accepted 30 August 2011.**

- Schadt, E. E. *et al.* Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**, 297–302 (2003).
- Morley, M. *et al.* Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**, 743–747 (2004).
- Stranger, B. E. *et al.* Genome-wide associations of gene expression variation in humans. *PLoS Genet.* **1**, e78 (2005).
- Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nature Genet.* **39**, 1494–1499 (2007).
- Heinzen, E. L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008).
- Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
- Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol. Psychiatry* **15**, 779–784 (2010).
- Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445–458 (2009).
- Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
- Somel, M. *et al.* MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* **20**, 1207–1218 (2010).
- Oldham, M. C. *et al.* Functional organization of the transcriptome in human brain. *Nature Neurosci.* **11**, 1271–1282 (2008).
- Torkamani, A., Dean, B., Schork, N. J. & Thomas, E. A. Coexpression network analysis of neural tissue reveals perturbations in developmental processes in schizophrenia. *Genome Res.* **20**, 403–412 (2010).
- Luo, L. & O'Leary, D. D. Axon retraction and degeneration in development and disease. *Annu. Rev. Neurosci.* **28**, 127–156 (2005).
- Deo, M., Yu, J. Y., Chung, K. H., Tippens, M. & Turner, D. L. Detection of mammalian microRNA expression by in situ hybridization with RNA oligonucleotides. *Dev. Dyn.* **235**, 2538–2548 (2006).
- Gao, F. B. Context-dependent functions of specific microRNAs in neuronal development. *Neural Develop.* **5**, 25 (2010).
- Coolen, M. & Bally-Cuif, L. MicroRNAs in brain development and physiology. *Curr. Opin. Neurobiol.* **19**, 461–470 (2009).
- Delaloy, C. *et al.* MicroRNA-9 coordinates proliferation and migration of human embryonic stem cell-derived neural progenitors. *Cell Stem Cell* **6**, 323–335 (2010).
- Loerch, P. M. *et al.* Evolution of the aging brain transcriptome and synaptic regulation. *PLoS ONE* **3**, e3329 (2008).
- Lu, T. *et al.* Gene regulation and DNA damage in the ageing human brain. *Nature* **429**, 883–891 (2004).
- Blalock, E. M. *et al.* Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc. Natl Acad. Sci. USA* **101**, 2173–2178 (2004).
- Colantuoni, C., Henry, G., Zeger, S. & Pevsner, J. SNOMAD (Standardization and Normalization of MicroArray Data): web-accessible gene expression data analysis. *Bioinformatics* **18**, 1540–1541 (2002).
- Leek, J. T. & Storey, J. D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* **3**, 1724–1735 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the families who donated tissue to make this study possible. We also thank the Offices of the Chief Medical Examiner of the District of Columbia, and of the Commonwealth of Virginia, Northern District, and the National Institute of Child and Health Development Brain and Tissue Bank for their collaboration. We thank R. McKay, N. Schork, F. McMahon and S. Zeger for their consultations on many issues, L. Marchionni for his assistance in assembling functional gene groups, as well as A. Deep-Soboslay, L. B. Bigelow, L. Wang, R. Buerlein, H. Choxi, V. Imamovic, Y. Snitkovsky, J. D. Paltan-Ortiz, J. Sirovatka, K. Becker, E. Lehrman and R. Vakkalanka for their contributions to this work.

**Author Contributions** C.C., design of the study, data exploration and analysis, writing of manuscript; B.K.L., design of the study, preparation of samples, data analysis, writing of the paper; T.Y., data analysis and web tool construction; T.M.H., brain collection, diagnosis, dissection (primary); writing/editing and commentary on analysis (secondary); planning experiment (primary); R.T., genotyping; J.T.L., surrogate variable analysis methods and code, statistical consultation; E.A.C., linear model methods, statistical consultation; A.G.E., microarray experiments; M.M.H., tissue characterization and micro/macro neuropathology; D.R.W., design and planning of the study, writing of manuscript; J.E.K., experimental design, characterization of specimens, data analysis and writing/editing.

**Author Information** The entire data set has been deposited in Gene Expression Omnibus under accession number GSE30272 and dbGaP under accession number phs000417.v1.p1 and can also be interrogated at <http://www.libd.org/braincloud>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.E.K. ([kleinmaj@mail.nih.gov](mailto:kleinmaj@mail.nih.gov)).

## METHODS

**Human post-mortem brain tissue collection.** The NIMH Brain Tissue Collection in the Clinical Brain Disorders Branch (NIMH, CBDB) obtained post-mortem human brains at autopsy primarily from the Offices of the Chief Medical Examiner of the District of Columbia, and of the Commonwealth of Virginia, Northern District, all with informed consent from the legal next of kin (protocol 90-M-0142 approved by the NIMH/NIH Institutional Review Board). Additional post-mortem fetal, infant, child and adolescent brain tissue samples were provided by the National Institute of Child Health and Human Development Brain and Tissue Bank for Developmental Disorders (<http://www.BTBank.org>) under contracts NO1-HD-4-3368 and NO1-HD-4-3383. The Institutional Review Board of the University of Maryland at Baltimore and the State of Maryland approved the protocol, and the tissue was donated to the NIMH under the terms of a Material Transfer Agreement. Clinical characterization, diagnoses, and macro- and microscopic neuropathological examinations were performed on all CBDB cases using a standardized paradigm. Details of tissue acquisition, handling, processing, dissection, clinical characterization, diagnoses, neuropathological examinations, RNA extraction and quality control measures were described previously<sup>23</sup>. The Brain and Tissue Bank cases were handled in a similar fashion (<http://medschool.umaryland.edu/BTBank/ProtocolMethods.html>). Toxicological analysis was performed on every case. Subjects with evidence of macro- or microscopic neuropathology, drug use, alcohol abuse, or psychiatric illness were excluded. Subject demographics and sample details are contained in Supplementary Table 7.

**RNA resources and analysis.** Post-mortem tissue homogenates of PFC grey matter (DLPFC, that is, BA46/9 in postnatal samples and the corresponding region of PFC in fetal samples) were obtained from all subjects ( $n = 269$  after all exclusion criteria). Total RNA was extracted from ~100 mg of tissue using the RNeasy kit (Qiagen) according to the manufacturer's protocol. Samples with RNA integrity number (RIN) <5 were excluded. 500 ng of each total RNA sample was reverse transcribed with an oligo dT-T7 and amplified (T7) using the Ambion MessageAmp II kit (catalogue no. 1753, Ambion). The generated aminoallyl UTP-labelled antisense RNAs (aRNAs) were then coupled with Cy3 mono NHS ester CyDye from GE Healthcare. Reference RNA was pooled from all samples and was treated identically to sample RNAs, but was labelled with the Cy5 fluorescent dye. Two-colour custom-spotted oligonucleotide microarrays from the NHGRI microarray core facility using the Illumina Oligoset (HEEBO7) of 49,152 70-mer probes were used. After purification, the labelled aRNAs are hybridized overnight to the oligo arrays in  $5\times$  SSC, 25% formamide and 0.2% SDS buffer at 45 °C using Maui Mixer FL hybridization chambers (BioMicro Systems). The slides are then washed at room temperature in a series of SSC/SDS buffers and dried by centrifugation. A laser confocal scanner (Agilent Technologies) was used to scan the hybridized microarrays. DeArray software (Scanalytics, Inc.) was used to export intensity data. Probes that were non-human, nonspecific (that is, mapped to >1 expressed sequence), incorrectly annotated, or probes containing polymorphisms with minor allele frequency > 0.01 according to HapMap in either YRI or CEU populations were removed from the analysis. Intensities below an empirically determined low intensity cutoff of 5.3 on the  $\log_2$  scale were dropped from the data. Probes with fewer than half of the fetal or postnatal data points remaining after this step were removed. Additionally, outliers defined as >6 mean average deviations from the age-appropriate linear fit were removed. The total number of probes remaining was 30,176. After background correction on the linear scale,  $\log_2$  ratios (sample/reference) were normalized across mean  $\log_2$  fluorescent intensities using loess correction<sup>21</sup>. Missing data in the gene expression data matrix were imputed at this stage to enable both SVA and PCA. After normalization,  $\log_2$  ratios were further adjusted to reduce the impact of known and unknown sources of systematic noise on gene expression measures using SVA<sup>22</sup>. Two surrogate variables were generated and used to adjust  $\log_2$  ratios in all subsequent linear models. Correlation between the naively created surrogate variables and known sources of noise were evident: SV1 + RIN:  $r = 0.37$ ,  $P = 4.7 \times 10^{-10}$ ; SV2 + ArrayBatch:  $r = 0.73$ ;  $P < 2 \times 10^{-16}$ . All of these microarray data analyses were conducted using custom code and tools from the Bioconductor project (<http://www.bioconductor.org/>) in the R statistical language (<http://www.r-project.org/>). Validation of microarray expression patterns was performed by Taqman qPCR (Supplementary Table 8).

**DNA resources and analysis.** DNA for genotyping was obtained from the cerebella of 266 of the total 269 samples in the collection and applied to either Illumina Infinium II 650K or Illumina Infinium HD Gemini 1M Duo BeadChips according to manufacturer's protocols. Only genotypes common to both platforms are

analysed here. Genotypes were called using BeadExpress software. SNPs were removed if the call rate was <98% (mean call rate for this study >99%), if not in Hardy-Weinberg equilibrium ( $P < 0.001$ ) within Caucasian and within African American races separately, or not polymorphic (MAF <0.01). The total number of SNPs remaining in the analysis was 625,439 (96.2%).

**Expression turning points.** For each probe, a linear-spline model of expression across age was fit with a single change point. The change point was allowed to vary across the entire age range, and the change point that produced the lowest mean-squared error was selected as the expression turning point for that probe. Data within the first and last decade of the range interrogated (0–10 and 70–80 years) were excluded to avoid edge effects ( $n = 7,272$  probes). See Fig. 1b (grey histogram).

**MDS using expression correlation as a distance metric.** This representation (see Fig. 1c) was generated using  $1 - r$  as a distance metric, where  $r$  is the pairwise Pearson's correlation coefficient calculated across all gene expression probes for each pair of samples. These distances were coupled with an MDS algorithm to attempt to satisfy all the pairwise distances in two-dimensional space. For both MDS and PCA, three-dimensional analyses more precisely depict systematic age effects (Supplementary Fig. 1, parts 2 and 3). It is important to note in this analysis (and those shown in Figs 1b and 4) that because expression data are expressed as a ratio to reference here, the mean expression correlation across all samples is near zero:  $r = 0.02$ .

**Functional gene groupings.** To generate functional gene groups for the analysis described in Fig. 2a and the text, we annotated all probes with data from Kyoto Encyclopedia of Genes and Genomes Pathways (<http://www.genome.jp/kegg/>), the Gene Ontology project (<http://www.geneontology.org>), the Pfam database (<http://www.sanger.ac.uk/Software/Pfam>), mouse knockout phenotypes and human disease phenotypes collected by Kevin Becker's group at the National Institute on Ageing<sup>24,25</sup>, the GSA project at Stanford (<http://www-stat.stanford.edu/~tibs/GSA>) and the GSEA project at the Broad Institute (<http://www.broad.mit.edu/gsea>), the HPRD project (<http://www.hprd.org>), as well as many groups collected from diverse sources at NCBI (<http://www.ncbi.nlm.nih.gov>), including protein-protein interactions and miRNA binding motifs<sup>26,27</sup>. Compilation of functional information from all of these sources and considering only gene groups of size 3–1,000 resulted in 23,810 partially redundant and overlapping functionally related gene groups. Enrichment of functional gene groups within various gene lists as described in the text was assessed by a standard hypergeometric test.

During the exploration of data for Fig. 2, it was observed that there is significant correlation between the age and sex variables within the first 6 months of life ( $r = 0.5$ ). To ensure that this correlation was not responsible for the discoveries in Fig. 2, the entire analysis was repeated while adjusting expression measures for sex. All the findings detailed in Fig. 2 were replicated in this verification analysis.

**Calculation of SNP-expression associations.** SNP-expression associations referred to in Fig. 3 and Supplementary Table 6 were carried out using linear models that included surrogate variables generated as described above, age, life stage (as defined in Fig. 1b), an interaction of age and life stage, sex, race, and the SNP under investigation. The SNP was included as a continuous variable, coded as 1,2,3; that is, an additive or 'dosage' model rather than a categorical or 'co-dominant' model was used. Association methods used to generate hits for the genetic distance used in Supplementary Fig. 2, part 1, were identical to those described above except for the omission of the race and sex terms in the linear model. To confirm that this analytical framework is capable of discovering associations between genetic and transcriptional metrics as analysed here, we conducted a positive control analysis (Supplementary Fig. 2, part 2; negative control also included).

23. Lipska, B. K. *et al.* Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biol. Psychiatry* **60**, 650–658 (2006).
24. Zhang, Y. *et al.* Systematic analysis, comparison, and integration of disease based human genetic association data and mouse genetic phenotypic information. *BMC Med. Genomics* **3**, 1 (2010).
25. De, S., Zhang, Y., Garner, J. R., Wang, S. A. & Becker, K. G. Disease and phenotype gene set analysis of disease based gene expression in mouse and human. *Physiol. Genomics* **42A**, 162–167 (2010).
26. Daniel, V. C. *et al.* A primary xenograft model of small-cell lung cancer reveals irreversible changes in gene expression imposed by culture *in vitro*. *Cancer Res.* **69**, 3364–3373 (2009).
27. Schaeffer, E. M. *et al.* Androgen-induced programs for prostate epithelial growth and invasion arise in embryogenesis and are reactivated in cancer. *Oncogene* **27**, 7180–7191 (2008).

# RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia

Johannes Zuber<sup>1,2\*</sup>, Junwei Shi<sup>1,3\*</sup>, Eric Wang<sup>1</sup>, Amy R. Rappaport<sup>1,4</sup>, Harald Herrmann<sup>5</sup>, Edward A. Sison<sup>6</sup>, Daniel Magoon<sup>6</sup>, Jun Qi<sup>7</sup>, Katharina Blatt<sup>8</sup>, Mark Wunderlich<sup>9</sup>, Meredith J. Taylor<sup>1</sup>, Christopher Johns<sup>1</sup>, Agustin Chicas<sup>1</sup>, James C. Mulloy<sup>9</sup>, Scott C. Kogan<sup>10</sup>, Patrick Brown<sup>6</sup>, Peter Valent<sup>5,8</sup>, James E. Bradner<sup>7</sup>, Scott W. Lowe<sup>1,4,11</sup> & Christopher R. Vakoc<sup>1</sup>

Epigenetic pathways can regulate gene expression by controlling and interpreting chromatin modifications. Cancer cells are characterized by altered epigenetic landscapes, and commonly exploit the chromatin regulatory machinery to enforce oncogenic gene expression programs<sup>1</sup>. Although chromatin alterations are, in principle, reversible and often amenable to drug intervention, the promise of targeting such pathways therapeutically has been limited by an incomplete understanding of cancer-specific dependencies on epigenetic regulators. Here we describe a non-biased approach to probe epigenetic vulnerabilities in acute myeloid leukaemia (AML), an aggressive haematopoietic malignancy that is often associated with aberrant chromatin states<sup>2</sup>. By screening a custom library of small hairpin RNAs (shRNAs) targeting known chromatin regulators in a genetically defined AML mouse model, we identify the protein bromodomain-containing 4 (Brd4) as being critically required for disease maintenance. Suppression of Brd4 using shRNAs or the small-molecule inhibitor JQ1 led to robust antileukaemic effects *in vitro* and *in vivo*, accompanied by terminal myeloid differentiation and elimination of leukaemia stem cells. Similar sensitivities were observed in a variety of human AML cell lines and primary patient samples, revealing that JQ1 has broad activity in diverse AML subtypes. The effects of Brd4 suppression are, at least in part, due to its role in sustaining Myc expression to promote aberrant self-renewal, which implicates JQ1 as a pharmacological means to suppress MYC in cancer. Our results establish small-molecule inhibition of Brd4 as a promising therapeutic strategy in AML and, potentially, other cancers, and highlight the utility of RNA interference (RNAi) screening for revealing epigenetic vulnerabilities that can be exploited for direct pharmacological intervention.

AML represents a paradigm for understanding how complex patterns of cooperating genetic and epigenetic alterations lead to tumorigenesis<sup>3,4</sup>. Although this complexity poses a challenge for the development of targeted therapies, diverse gene mutations in AML generally converge functionally to deregulate similar core cellular processes. One key event in AML initiation is the corruption of cell-fate programs to generate leukaemia stem cells that aberrantly self-renew and thereby maintain and propagate the disease<sup>5</sup>. Although it is incompletely understood, this process has been linked to changes in regulatory chromatin modifications<sup>2</sup>. For example, common AML oncogenes such as those encoding the AML1-ETO and MLL fusion proteins induce self-renewal programs at least in part through reprogramming of epigenetic pathways<sup>6,7</sup>. In addition, several genes encoding epigenetic regulators have been identified as targets of somatic mutation in AML<sup>8,9</sup>. Because epigenetic

alterations induced by oncogenic stimuli are potentially reversible, chromatin regulators are being explored as candidate drug targets<sup>1</sup>.

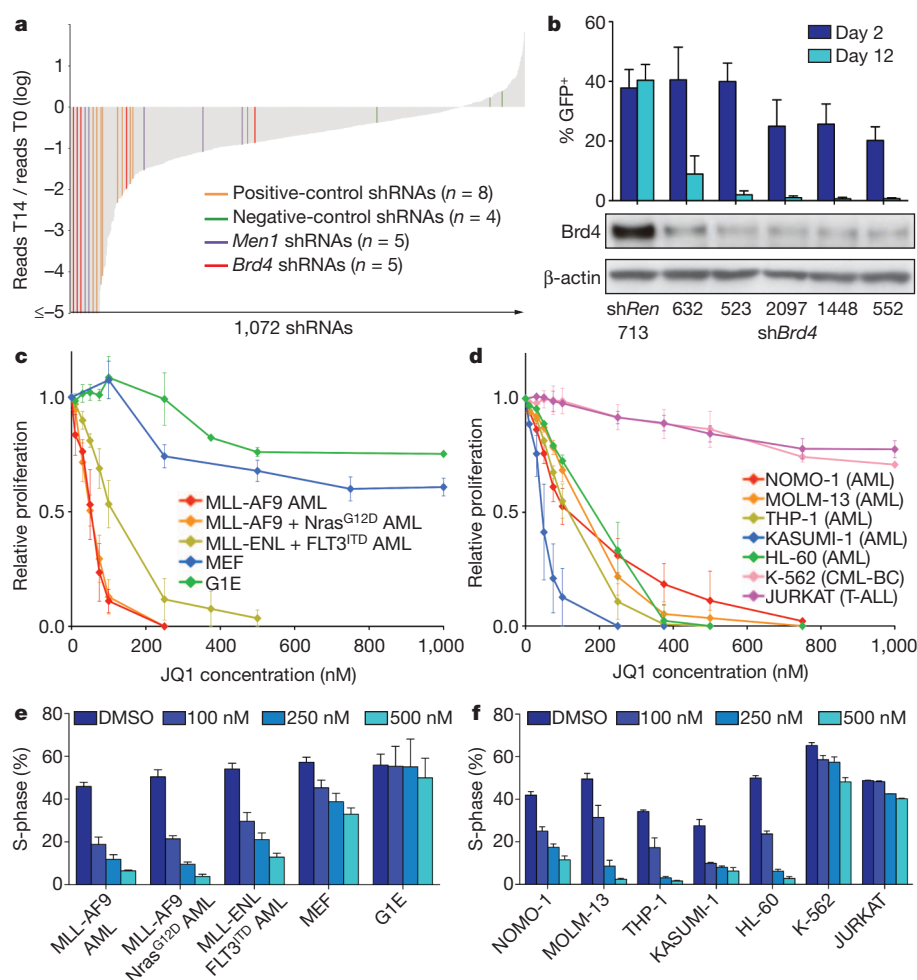
To probe epigenetic pathways required for AML maintenance systematically, we built a custom shRNA library targeting 243 known chromatin regulators, including most 'writers', 'readers' and 'erasers' of epigenetic marks (Supplementary Fig. 1 and Supplementary Table 1). This library of 1,094 shRNAs (3–6 per gene) was constructed in TRMPV-Neo, a vector optimized for negative-selection RNAi screening, and was transduced as one pool into an established Tet-on-competent AML mouse model driven by MLL-AF9 and Nras<sup>G12D</sup> (ref. 10). After drug selection, shRNA expression was induced by addition of doxycycline, and changes in library representation after 14 days of culture were monitored using deep sequencing of shRNA guide strands amplified from genomic DNA (Fig. 1a and Supplementary Fig. 2). Using the scoring criterion of more than twenty-fold depletion in each of two independent replicates, 177 shRNAs were strongly depleted. These included all eight positive-control shRNAs targeting essential genes (*Rpa1*, *Rpa3*, *Pcna* and *Polr2b*), as well as several shRNAs targeting two known MLL-AF9 cofactors (*Men1* and *Psip1*)<sup>11,12</sup>. Genes for which at least two independent shRNAs scored were subjected to extensive one-by-one validation using an independent MLL-AF9/Nras<sup>G12D</sup> AML cell line and vector system (Supplementary Fig. 3a). In both the primary screen and validation stages, several shRNAs targeting *Brd4* were among the most strongly depleted, identifying this gene as the top scorer in the screen (Fig. 1a and Supplementary Fig. 3b).

Brd4 is a member of the BET family of bromodomain-containing proteins that bind to acetylated histones to influence transcription<sup>13</sup>. *BRD4* is also a proto-oncogene that can be mutated via chromosomal translocation in a rare form of squamous-cell carcinoma<sup>14</sup>, although a role in leukaemia has not been described. The recent development of small-molecule BET bromodomain inhibitors<sup>15,16</sup>, together with our screening results, prompted us to investigate the suitability of Brd4 as an AML drug target. Five independent *Brd4* shRNAs showed a close correspondence between knockdown efficiency and growth inhibition, indicating on-target effects (Fig. 1b). Suppression of Brd4 led to cell-cycle arrest and apoptosis of leukaemia cells, whereas the equivalent knockdown in immortalized murine embryonic fibroblasts (MEFs) led to only modest cell-cycle inhibition without cytotoxicity (Supplementary Fig. 4a–d). Brd4 knockdown also failed to influence the growth of non-transformed G1E erythroblast cells (Supplementary Fig. 4e). In addition, shRNAs targeting *BRD4* were sufficient to induce cell-cycle arrest in two MLL-AF9<sup>+</sup> human AML lines (Supplementary Fig. 5). Together, these results indicate that Brd4 is a critical requirement in MLL-AF9-induced AML.

<sup>1</sup>Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. <sup>2</sup>Research Institute of Molecular Pathology (IMP), Dr. Bohr-Gasse 7, A-1030 Vienna, Austria. <sup>3</sup>Molecular and Cellular Biology Program, Stony Brook University, Stony Brook, New York 11794, USA. <sup>4</sup>Watson School of Biological Sciences, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. <sup>5</sup>Ludwig Boltzmann Cluster Oncology, Medical University of Vienna, A-1090 Vienna, Austria. <sup>6</sup>Division of Pediatric Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21231, USA.

<sup>7</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. <sup>8</sup>Department of Internal Medicine I, Division of Hematology & Hemostaseology, Medical University of Vienna, A-1090 Vienna, Austria. <sup>9</sup>Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. <sup>10</sup>Department of Laboratory Medicine, University of California at San Francisco, San Francisco, California 94143, USA. <sup>11</sup>Howard Hughes Medical Institute, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA.

\*These authors contributed equally to this work.



**Figure 1 | AML growth is sensitive to Brd4 inhibition.** **a**, Pooled negative-selection screening in MLL-AF9/Nras<sup>G12D</sup> leukaemia, depicting changes in representation of 1,072 informative shRNAs during 14 days of culture. shRNA abundance ratios were calculated as the number of reads after 14 days of culture on doxycycline (T14) divided by the number of reads before doxycycline treatment (T0), and plotted as the mean of two replicates in ascending order. Completely depleted shRNAs (0 reads at T14,  $n = 71$ ) were plotted as a ratio of  $10^{-5}$ ; highlighted shRNAs in this group are shown with even spacing in alphabetical order. Positive controls included shRNAs targeting *Rpa1*, *Rpa3*, *Pcna* and *Polr2b*. Negative-control shRNAs targeted renilla luciferase or *Braf*. **b**, Percentage of GFP-positive MLL-AF9/Nras<sup>G12D</sup> leukaemia cells 2 days and 12 days after transduction with LMN constructs expressing the indicated shRNAs (top panel). Western blotting of whole-cell lysates prepared from MEF cultures transduced with the indicated TiTMPV-Neo-shRNAs and induced with doxycycline for 5 days (bottom panel). shRen, renilla luciferase shRNA. Representative experiments are shown. **c, d**, Proliferation rates of JQ1-treated cells, calculated by measuring the increase in viable cell number after 72 h in culture and fitting data to an exponential growth curve. Results are normalized to the proliferation rate of vehicle/DMSO-treated cells, set to 1 ( $n = 3$ ). CML-BC, chronic myeloid leukaemia blast crisis; T-ALL, T-cell acute lymphoblastic leukaemia. **e, f**, Percentage of cells in S-phase (bromodeoxyuridine (BrdU)<sup>+</sup>) after JQ1 treatment for 48 h at the indicated concentrations ( $n = 3$ ). BrdU was pulsed for 30 min in all experiments shown. All error bars represent s.e.m.

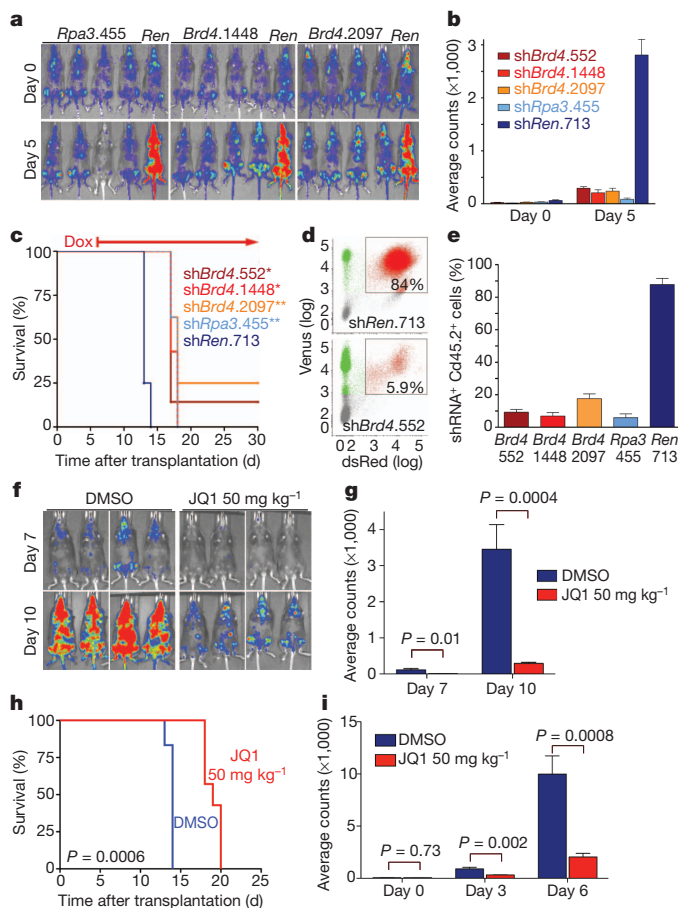
Next, we examined the sensitivity of leukaemia cells to JQ1, a first-in-class small-molecule inhibitor of BET bromodomains with highest affinity for the first bromodomain of Brd4 (ref. 15). Proliferation of mouse MLL-fusion leukaemia cells was notably sensitive to sub-micromolar JQ1 concentrations, as compared to proliferation of fibroblasts and G1E cells (Fig. 1c), correlating with the relative impact of *Brd4* shRNAs on proliferation of these different cell types. We also examined growth-inhibitory effects of JQ1 in a series of established human leukaemia cell lines as well as in adult and paediatric primary leukaemia samples. We observed broad growth-suppressive activity of JQ1 ( $IC_{50} < 500$  nM) in 13 out of 14 AML cell lines (Fig. 1d and Supplementary Fig. 6a, b) and in 12 out of 15 primary human AML samples, representing diverse disease subtypes (Supplementary Fig. 7, 8). In addition, all three primary MLL-rearranged infant leukaemias tested were sensitive to JQ1 (Supplementary Fig. 8), whereas other non-AML-leukaemia and solid-tumour cell lines showed minimal sensitivity to the compound (Fig. 1d and Supplementary Fig. 6c). In all AML lines examined, JQ1 treatment triggered cell-cycle arrest and apoptosis, similar to the effects seen upon shRNA-mediated Brd4 knockdown (Fig. 1e, f and Supplementary Figs 7–9). Together, these data indicate a critical requirement for Brd4 in AML proliferation that can be effectively inhibited using the bromodomain inhibitor JQ1.

We next investigated the relevance of Brd4 to AML progression *in vivo*. To suppress Brd4 in established AML in mice, Tet-on-competent MLL-AF9/Nras<sup>G12D</sup> leukaemia cells were transduced with TRMPV-Neo constructs containing *Brd4* shRNAs or control shRNAs, and transplanted into secondary recipient mice (Supplementary Fig. 10). After disease onset, confirmed by bioluminescent imaging, shRNA expression was induced by doxycycline administration. Subsequent monitoring showed that *Brd4* knockdown resulted in a marked delay

in leukaemia progression and a survival benefit (Fig. 2a–c and Supplementary Fig. 11). Taking advantage of the dsRed reporter linked to shRNA expression in the TRMPV-Neo vector<sup>10</sup>, flow-cytometry analysis verified that *Brd4*-shRNA-positive cells were depleted in the terminal leukaemia burden as compared to controls, indicating that the mice succumbed to an outgrowth of *Brd4*-shRNA-negative cells (Fig. 2d, e). Together, these data indicate that RNAi-mediated suppression of Brd4 inhibits leukaemia progression *in vivo*.

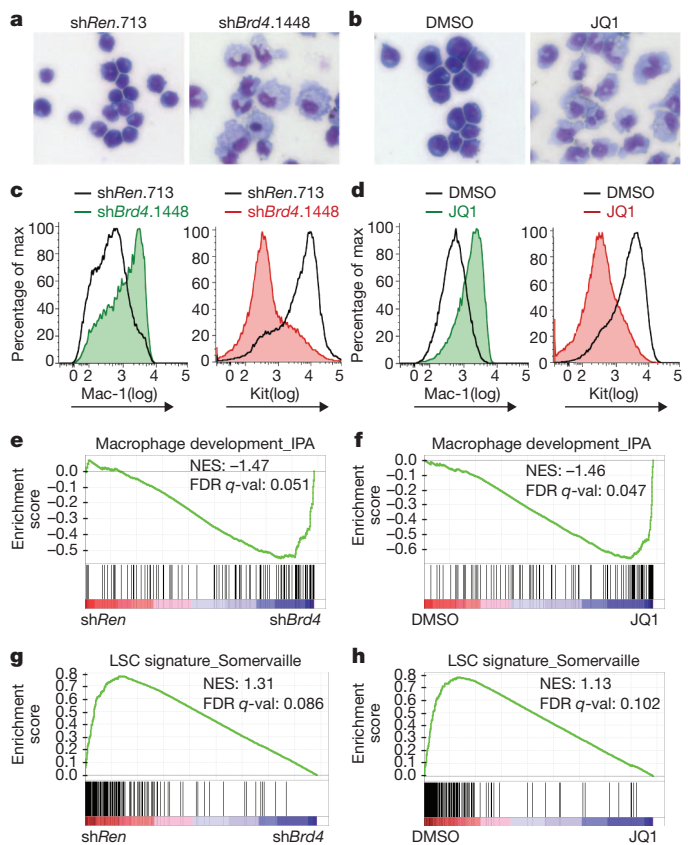
To examine whether JQ1 has single-agent activity in AML, mice transplanted with MLL-AF9/Nras<sup>G12D</sup> leukaemia cells were treated either with daily injections of JQ1 (50 mg kg<sup>-1</sup>) or with vehicle. JQ1 administration led to a marked delay in disease progression and significantly extended survival (Fig. 2f–h). JQ1 also showed single-agent activity in an intervention setting, in which treatment was initiated only after disease was detected by bioluminescent imaging (Fig. 2i and Supplementary Fig. 12). Comparable effects were observed in an independent AML mouse model based on expression of AML1-ETO9a and Nras<sup>G12D</sup> and loss of p53 (ref. 17), which is known to be insensitive to conventional chemotherapy (Supplementary Fig. 13). Consistent with previous findings<sup>15</sup>, JQ1 treatment was well tolerated in mice, with little if any impact on normal haematopoiesis (Supplementary Figs 14–16). Collectively, these findings demonstrate that JQ1 has potent and leukaemia-specific effects as a single agent *in vivo*.

AML is characterized by an expanded self-renewal capacity linked with an inability to complete terminal myeloid differentiation. Therefore, we next considered whether Brd4 influences the differentiation state of leukaemia cells. Both expression of *Brd4* shRNA and JQ1 treatment altered the morphology of MLL-AF9/Nras<sup>G12D</sup> leukaemia from myelomonocytic blasts to cells with a macrophage-like appearance (Fig. 3a, b and Supplementary Fig. 17a). Upon Brd4 inhibition,



**Figure 2 | Brd4 is required for AML progression *in vivo*.** **a**, Bioluminescent imaging of mice transplanted with MLL-AF9/Nras<sup>G12D</sup> leukaemia cells harbouring the indicated TRMPV-Neo-shRNAs. Doxycycline was administered upon disease onset, 6 days after transplant. Day 0 indicates the time of doxycycline treatment. **b**, Quantification of bioluminescent imaging responses after doxycycline treatment. Mean values of four replicate mice are shown. **c**, Kaplan-Meier survival curves of recipient mice transplanted with the indicated TRMPV-Neo-shRNA leukaemia lines. The interval of doxycycline treatment is indicated by the arrow. Each shRNA group contained 6–8 mice. Statistical significance compared to shRen was calculated using a log-rank test; \*,  $P = 0.0001$ ; \*\*,  $P < 0.0001$ . **d**, Representative flow cytometry plots of donor-derived (Cd45.2<sup>+</sup>) bone marrow cells in terminally diseased doxycycline-treated mice. The gate shown includes dsRed<sup>+</sup>/shRNA<sup>+</sup> cells. **e**, Percentage of dsRed<sup>+</sup>/shRNA<sup>+</sup> cells in the Cd45.2<sup>+</sup> terminal leukaemia burden. Mean values of four replicate mice are shown. **f**, Bioluminescent imaging of MLL-AF9/Nras<sup>G12D</sup> leukaemia recipient mice at the indicated day after initiation of treatment with JQ1 (50 mg kg<sup>-1</sup> d<sup>-1</sup>) or DMSO carrier. **g**, Quantification of bioluminescent imaging responses to JQ1 treatment. Mean values of six DMSO- and seven JQ1-treated mice are shown.  $P$  values were calculated using a two-tailed Student's  $t$ -test. **h**, Kaplan-Meier survival curves of control and JQ1-treated mice. Statistical significance was calculated using a log-rank test. In **f**, **g** and **h**, JQ1 treatment was initiated on day 1 after transplant of 50,000 leukaemia cells. **i**, Quantification of bioluminescent imaging responses to JQ1 treatment in established disease. Treatment of leukaemia mice was initiated 6 days after transplant, when disease first became detectable by imaging. Mean values of six DMSO- and seven JQ1-treated mice are shown.  $P$  values were calculated using a two-tailed Student's  $t$ -test. All error bars represent s.e.m.

leukaemia cells showed increased surface expression of integrin  $\alpha$ M (Itgam, also known as Mac-1), a myeloid differentiation marker, and decreased expression of Kit, a marker associated with leukaemia stem cells (LSCs) in mouse models of MLL-rearranged leukaemia (Fig. 3c, d and Supplementary Fig. 17b, c)<sup>18,19</sup>. In addition, JQ1 treatment induced morphological signs of maturation phenotypes in most of the primary leukaemia samples tested, albeit to varying degrees (Supplementary Figs 7 and 8).



**Figure 3 | Brd4 inhibition leads to myeloid differentiation and leukaemia stem-cell depletion.** **a**, **b**, Light microscopy of May-Grünwald/Giemsa-stained MLL-AF9/Nras<sup>G12D</sup> leukaemia cells after 2 days of doxycycline-induced shRNA expression or 2 days of JQ1 treatment (100 nM). Expression of shRNA was induced in TRMPV-Neo-transduced leukaemia cells. Imaging was performed with a  $\times 40$  objective. Representative images of three biological replicates are shown. **c**, **d**, Flow cytometry analysis of Mac-1 and Kit surface expression after 4 days of shRNA expression or 2 days of JQ1 treatment (100 nM). A representative experiment of three biological replicates is shown. **e**–**h**, GSEA plots evaluating changes in macrophage and LSC gene signatures upon Brd4 inhibition. In **e** and **g**, RNA for expression arrays was obtained from sorted dsRed<sup>+</sup>/shRNA<sup>+</sup> cells (shRen versus three different Brd4 shRNAs) after 2 days of doxycycline induction. In **f** and **h**, microarray data were obtained from leukaemia cells treated for 2 days with DMSO or 100 nM JQ1. NES, normalized enrichment score; FDR  $q$ -val, false discovery rate  $q$ -value (the probability that a gene set with a given NES represents a false-positive finding).

To further investigate whether suppression of Brd4 affects the LSC compartment, we performed gene set enrichment analysis (GSEA) of expression microarray data obtained from Brd4-shRNA-treated and JQ1-treated leukaemia cells<sup>20</sup>. GSEA revealed a marked upregulation of macrophage-specific gene expression after Brd4 inhibition (Fig. 3e, f), as well as global downregulation of a gene expression signature previously shown to discriminate LSCs from non-self-renewing leukaemia cell populations (Fig. 3g, h)<sup>21</sup>. A similar profile of gene expression changes was seen after JQ1 treatment of THP-1 cells, a human MLL-AF9-expressing AML cell line (Supplementary Fig. 18). Although we cannot exclude the involvement of additional cellular targets, the strong concordance between phenotypes induced by Brd4 shRNAs and JQ1 supports Brd4 as the relevant target of JQ1 in AML. Together, these findings indicate that Brd4 is critically required to maintain LSC populations and prevent terminal myeloid differentiation.

Recent evidence indicates that the Myc transcriptional network has an important role in LSC self-renewal<sup>22,23</sup>. Because previous studies also implicate Myc as a potential downstream target of Brd4 (refs 24, 25), we examined whether this regulatory function was relevant to the anti-leukaemic effects of JQ1. Brd4 inhibition with shRNAs or JQ1 led to a

marked reduction of mRNA and protein levels of Myc in MLL-AF9/Nras<sup>G12D</sup> leukaemia, whereas these effects were minimal in MEF and G1E cells (Fig. 4a, b and Supplementary Fig. 19a–c). Downregulation of Myc mRNA levels occurred within 60 min of JQ1 exposure, qualitatively preceding the increased expression of genes related to macrophage differentiation, such as *Cd74* (Fig. 4c). Supporting a direct role in Myc transcriptional regulation, chromatin immunoprecipitation (ChIP) experiments identified a region of focal Brd4 occupancy about 2 kilobases upstream of the *Myc* promoter and this was eliminated after exposure to competitive JQ1 (Fig. 4d). As expected, RNAi- or JQ1-induced suppression of Brd4 led to a global reduction in expression of

Myc target genes<sup>22,26</sup> (Supplementary Figs 19d and 20). Notably, JQ1 treatment triggered Myc downregulation in a broad array of mouse and human leukaemia cell lines examined (Fig. 4a, b and Supplementary Fig. 21), indicating that JQ1 may provide a means to suppress the Myc pathway in a range of AML subtypes.

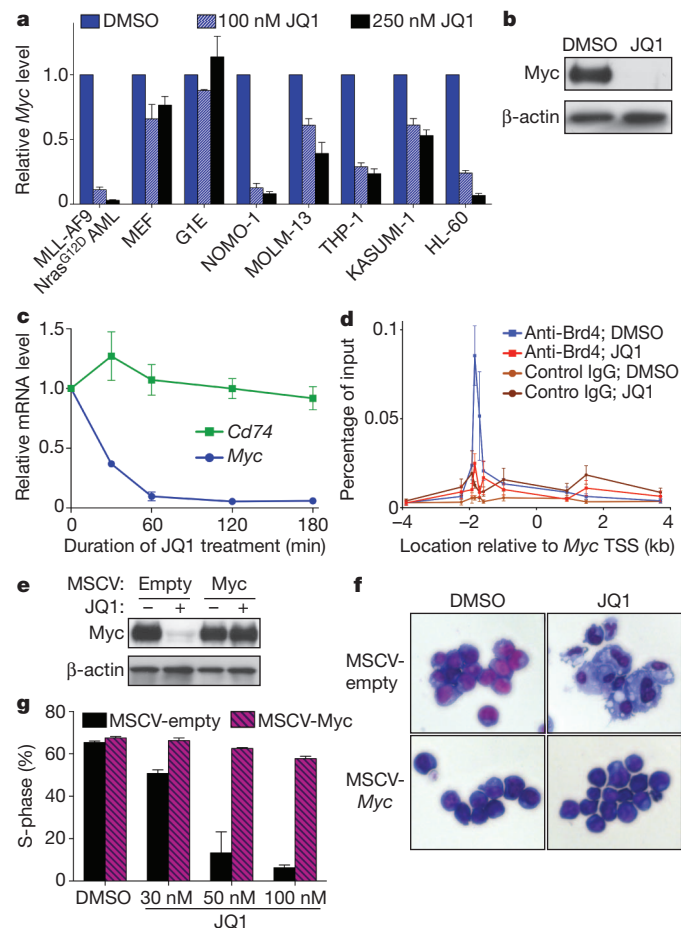
To evaluate whether suppression of Myc confers the growth-inhibitory effects of JQ1, we generated MLL-AF9/Nras<sup>G12D</sup> leukaemia cultures in which the *Myc* cDNA was ectopically expressed from a retroviral promoter, which resulted in *Myc* expression levels that were only slightly elevated but entirely resistant to JQ1-induced suppression (Fig. 4e and Supplementary Fig. 22a). Ectopic *Myc* expression conferred nearly complete resistance to macrophage differentiation and cell-cycle arrest induced by JQ1 and *Brd4* shRNAs (Fig. 4f, g and Supplementary Figs 22b and 23). Furthermore, global gene-expression profiling revealed that most of the JQ1-induced transcriptional changes are probably secondary effects of Myc downregulation (Supplementary Fig. 24). Ectopic *Myc* expression was unable to prevent JQ1-induced cell death, indicating that Brd4 has Myc-independent roles in regulating cell survival (Supplementary Fig. 22c, d). These findings collectively support a role for Brd4 in maintaining Myc expression to preserve an undifferentiated cellular state in AML.

By taking a non-biased screening approach targeting epigenetic regulators, our study has identified Brd4 as a critical factor required for AML disease maintenance. Because Brd4 is not evidently mutated or overexpressed in AML (Supplementary Fig. 25), the exquisite sensitivity of leukaemia cells to Brd4 inhibition would not have been revealed simply through genetic or transcriptional characterization of this disease. We further show that the bromodomain inhibitor JQ1 has broad activity in diverse AML contexts and, by comparing its effects to those induced by *Brd4* shRNAs, we provide evidence that Brd4 is the primary target for the antileukaemic activity of JQ1. Of note, JQ1 is a first-generation chemical inhibitor yet to be optimized for *in vivo* delivery, with a half-life of only about 1 h in rodents (Supplementary Fig. 26 and ref. 15). The more robust antileukaemic effects seen using *Brd4* shRNAs *in vivo* indicate that second-generation derivatives of this compound may have greater clinical activity. Regardless, our results unambiguously highlight the utility of RNAi screening for revealing candidate drug targets in cancer.

As a competitive inhibitor of the acetyl-lysine binding domain, JQ1 interferes with the ability of Brd4 to 'read' histone acetylation marks that facilitate transcriptional activation<sup>15</sup>. When applied to leukaemia cells, JQ1 interferes with transcriptional circuits supporting self-renewal, thus targeting LSCs and inducing terminal differentiation. In a parallel study, we identified the transcription factor Myb as a critical mediator of addition to the MLL-AF9 oncogene<sup>31</sup>. Notably, global gene-expression changes observed after genetic or pharmacological inhibition of Brd4 are remarkably similar to those seen upon suppressing Myb<sup>31</sup>, indicating that Myb and Brd4 may intersect functionally in a common transcriptional circuit that is essential for malignant self-renewal. A key downstream effector of both Myb and Brd4 is the oncoprotein Myc (ref. 27), which has been validated as an attractive therapeutic target but has thus far escaped efforts at pharmacological inhibition<sup>28,29</sup>. Although the precise mechanism remains to be further defined, targeting Brd4 abolishes Myc expression and limits self-renewal, with selectivity for the malignant context, thus avoiding the haematopoietic toxicities that may be associated with systemic Myc inhibition<sup>30</sup>. As such, our study may define a general strategy to disarm oncogenic pathways through the direct modulation of the epigenetic machinery.

## METHODS SUMMARY

**Pooled negative-selection RNAi screening.** A customized shRNA library targeting 243 chromatin-regulating mouse genes was designed using miR30-adapted BIOPREDsi predictions, and was generated by PCR-cloning a pool of oligonucleotides synthesized on 55k arrays (Agilent Technologies). Pools of shRNAs were subcloned into the TRMPV-Neo vector (Addgene catalogue no. 27990) together with control shRNAs, and transduced into Tet-on MLL-AF9/Nras<sup>G12D</sup> leukaemia cells for negative-selection screening, essentially as described previously<sup>10</sup>. All



**Figure 4 | JQ1 suppresses the Myc pathway in leukaemia cells.** **a**, Quantitative reverse transcription PCR (qRT-PCR) of relative *Myc* mRNA levels in the indicated mouse or human cells after a 48 h treatment with indicated JQ1 dose or DMSO. Results were normalized to *Gapdh*, with the relative mRNA level in untreated cells set to 1 ( $n = 3$ ). **b**, Western blotting of whole-cell lysates prepared from MLL-AF9/Nras<sup>G12D</sup> leukaemia cells treated for 48 h with DMSO or 250 nM JQ1. A representative experiment of three biological replicates is shown. **c**, qRT-PCR time course at indicated time points after treatment of MLL-AF9/Nras<sup>G12D</sup> leukaemia cells with 250 nM JQ1. Results were normalized to *Gapdh*, with the relative mRNA level in untreated cells set to 1 ( $n = 3$ ). **d**, ChIP-qPCR performed in MLL-AF9/Nras<sup>G12D</sup> leukaemia cells with the indicated antibodies and PCR primer locations ( $n = 6$  for DMSO;  $n = 4$  for JQ1-treated). TSS, transcription start site. **e**, Western blotting of whole-cell lysates prepared from MLL-AF9/Nras<sup>G12D</sup> leukaemia cells transduced with empty vector or a *Myc*-cDNA-containing MSCV retrovirus. Cells were treated for 48 h with DMSO or 250 nM JQ1. A representative experiment of three biological replicates is shown. **f**, Light microscopy of May-Grünwald/Giemsa-stained MLL-AF9/Nras<sup>G12D</sup> leukaemia cells transduced with an empty vector or with the *Myc* cDNA. Cells were treated for 5 days with 50 nM JQ1 and imaged using a  $\times 40$  objective. A representative image of three biological replicates is shown. **g**, Quantification of BrdU incorporation after a 30-min pulse in MLL-AF9/Nras<sup>G12D</sup> leukaemia cells transduced with empty control vector or the *Myc* cDNA. Cells were treated with JQ1 for 5 days at the indicated concentrations ( $n = 3$ ). All error bars shown represent s.e.m.

shRNA sequences as well as primary screening data are provided in Supplementary Table 1.

**Animal studies.** All mouse experiments were approved by the Cold Spring Harbor animal care and use committee. For conditional RNAi experiments *in vivo*, Tet-on MLL-AF9/Nras<sup>G12D</sup> leukaemia cells were transduced with TRMPV-Neo-shRNA constructs, followed by transplantation into sub-lethally irradiated recipient mice, as described previously<sup>10</sup>. For shRNA induction, animals were treated with doxycycline in both drinking water (2 mg ml<sup>-1</sup> with 2% sucrose; Sigma-Aldrich) and food (625 mg kg<sup>-1</sup>, Harlan laboratories). For JQ1 treatment trials, a stock of 100 mg ml<sup>-1</sup> JQ1 in DMSO was diluted 20-fold by dropwise addition of a 10% 2-hydroxypropyl- $\beta$ -cyclodextrin carrier (Sigma) under vortexing, yielding a 5 mg ml<sup>-1</sup> final solution. Mice were intraperitoneally injected daily with freshly diluted JQ1 (50 or 100 mg kg<sup>-1</sup>) or a similar volume of carrier containing 5% DMSO.

**Microarray analysis.** Expression microarrays were performed using Affymetrix ST 1.0 GeneChips. Raw microarray data can be accessed at Gene Expression Omnibus, GSE29799. Pathway analysis was performed using GSEA v2.07 software with 1,000 phenotype permutations<sup>20</sup>. All gene sets used for GSEA are provided in Supplementary Table 2.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 2 March; accepted 23 June 2011.**

**Published online 3 August 2011.**

- Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. *Nature Rev. Cancer* **10**, 457–469 (2010).
- Chen, J., Odenike, O. & Rowley, J. D. Leukaemogenesis: more than mutant genes. *Nature Rev. Cancer* **10**, 23–36 (2010).
- Gilliland, D. G., Jordan, C. T. & Felix, C. A. The molecular basis of leukemia. *Hematology (Am. Soc. Hematol. Educ. Program)* 80–97 (2004).
- Figuerola, M. E. *et al.* DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* **17**, 13–27 (2010).
- Dick, J. E. Stem cell concepts renew cancer research. *Blood* **112**, 4793–4807 (2008).
- Wang, J., Hoshino, T., Redner, R. L., Kajigaya, S. & Liu, J. M. ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. *Proc. Natl Acad. Sci. USA* **95**, 10860–10865 (1998).
- Krivtsov, A. V. *et al.* H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* **14**, 355–368 (2008).
- Delhommeau, F. *et al.* Mutation in *TET2* in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
- Ley, T. J. *et al.* *DNMT3A* mutations in acute myeloid leukemia. *N. Engl. J. Med.* **363**, 2424–2433 (2010).
- Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nature Biotechnol.* **29**, 79–83 (2011).
- Yokoyama, A. & Cleary, M. L. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* **14**, 36–46 (2008).
- Yokoyama, A. *et al.* The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123**, 207–218 (2005).
- Wu, S. Y. & Chiang, C. M. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J. Biol. Chem.* **282**, 13141–13145 (2007).
- French, C. A. *et al.* *BRD4-NUT* fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res.* **63**, 304–307 (2003).
- Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
- Nicodeme, E. *et al.* Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119–1123 (2010).
- Zuber, J. *et al.* Mouse models of human AML accurately predict chemotherapy response. *Genes Dev.* **23**, 877–889 (2009).
- Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10**, 257–268 (2006).
- Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818–822 (2006).
- Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
- Somervaille, T. C. *et al.* Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. *Cell Stem Cell* **4**, 129–140 (2009).
- Kim, J. *et al.* A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* **143**, 313–324 (2010).
- Wong, P. *et al.* The *miR-17-92* microRNA polycistron regulates MLL leukemia stem cell potential by modulating p21 expression. *Cancer Res.* **70**, 3833–3842 (2010).
- Jang, M. K. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell* **19**, 523–534 (2005).
- Yang, Z., He, N. & Zhou, Q. Brd4 recruits P-TEFb to chromosomes at late mitosis to promote G1 gene expression and cell cycle progression. *Mol. Cell. Biol.* **28**, 967–976 (2008).
- Schuhmacher, M. *et al.* The transcriptional program of a human B cell line in response to Myc. *Nucleic Acids Res.* **29**, 397–406 (2001).
- Schmidt, M., Nazarov, V., Stevens, L., Watson, R. & Wolff, L. Regulation of the resident chromosomal copy of *c-myc* by *c-Myb* is involved in myeloid leukemogenesis. *Mol. Cell. Biol.* **20**, 1970–1981 (2000).
- Soucek, L. *et al.* Modelling Myc inhibition as a cancer therapy. *Nature* **455**, 679–683 (2008).
- Felsher, D. W. & Bishop, J. M. Reversible tumorigenesis by *MYC* in hematopoietic lineages. *Mol. Cell* **4**, 199–207 (1999).
- Wilson, A. *et al.* c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev.* **18**, 2747–2763 (2004).
- Zuber, J. *et al.* An integrated approach to dissecting oncogene addiction implicates a Myb coordinated self-renewal program as essential for leukemia maintenance. *Genes Dev.* doi:10.1101/gad.1726911 (in the press).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank B. Ma, S. Muller and M. Weissenboeck for technical assistance; J. Simon, E. Earl and L. Bianco for support with mouse work; C. dos Santos for assistance with LSK FACS analysis; S. Hearn for microscopy support; G. Hannon, K. Chang, and E. Hodges for shRNA technology support; A. Gordon and M. Hammell for bioinformatics support; L. Dow for assistance with mouse pathology sample preparation; and G. Blobel for comments on the manuscript. We thank the Don Monti Memorial Research Foundation and Laurie Strauss Leukemia Foundation for research support. J.Z. was supported by a research fellowship from the German Research Foundation (DFG) and by the Andrew Seligson Memorial Clinical Fellowship at CSHL; A.R.R. was supported by an NIH traineeship and the Barbara McClintock fellowship; J.E.B. and J.Q. are supported by the Damon-Runyon Cancer Research Foundation and Smith Family Foundation. P.B. is supported by a Damon Runyon-Lilly Clinical Investigator Award and a Leukemia and Lymphoma Society (LLS) Translational Research Program Grant, and is an LLS Scholar in Clinical Research. S.W.L. is supported by a Specialized Center of Research (SCOR) grant from the Leukemia and Lymphoma Society of America, a Cancer Target Discovery and Development (CTD2) grant from the National Cancer Institute, and by the Howard Hughes Medical Institute; C.R.V., J.S. and E.W. are supported by the CSHL President's Council and the SASS Foundation for Medical Research.

**Author Contributions** J.Z., J.S. and C.R.V. designed research, performed experiments and analysed data. E.W., A.R.R. and M.J.T. performed experiments and analysed data. C.J. performed microarray experiments. A.C. assisted with shRNA library design. H.H., E.A.S., D.M., K.B., P.B. and P.V. performed experiments with primary leukaemia specimens. J.C.M. and M.W. generated and provided engineered AML lines. S.C.K. performed histological analysis. J.Q. and J.E.B. designed research and synthesized and supplied JQ1. J.Z., S.W.L. and C.R.V. co-wrote the paper. S.W.L. and C.R.V. supervised the research.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Raw microarray data can be accessed at Gene Expression Omnibus, GSE29799. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to C.R.V. ([vakoc@cshl.edu](mailto:vakoc@cshl.edu)) or S.W.L. ([lowe@cshl.edu](mailto:lowe@cshl.edu)).

## METHODS

**Pooled negative-selection RNAi screening.** A custom shRNA library targeting 243 chromatin-regulating mouse genes was designed using miR30-adapted BIOPREDSi predictions (six shRNAs per gene) and constructed by PCR-cloning a pool of oligonucleotides synthesized on 55k customized arrays (Agilent Technologies) as previously described<sup>10</sup>. After sequence verification, 1,094 shRNAs (3–6 per gene) were combined with several positive- and negative-control shRNAs at equal concentrations in one pool. This pool was subcloned into TRMPV-Neo and transduced into Tet-on MLL-AF9/Nras<sup>G12D</sup> leukaemia cells using conditions that predominantly lead to a single retroviral integration and represent each shRNA in a calculated number of >500 cells (a total of thirty million cells at infection, 2% transduction efficiency). Transduced cells were selected for 5 days using 1 mg ml<sup>-1</sup> G418 (Invitrogen); at each passage more than twenty million cells were maintained to preserve library representation throughout the experiment. After drug selection, T0 samples were obtained (~twenty million cells per replicate) and cells were subsequently cultured with 0.5 mg ml<sup>-1</sup> G418 and 1 µg ml<sup>-1</sup> doxycycline to induce shRNA expression. After 14 days (12 passages, T14), about fifteen million shRNA-expressing (dsRed<sup>+</sup>/Venus<sup>+</sup>) cells were sorted for each replicate using a FACSAriaII (BD Biosciences). Genomic DNA from T0 and T14 samples was isolated by two rounds of phenol extraction using PhaseLock tubes (5prime) followed by isopropanol precipitation. Deep-sequencing template libraries were generated by PCR amplification of shRNA guide strands as previously described<sup>10</sup>. Libraries were analysed on an Illumina Genome Analyser at a final concentration of 8 pM; 18 nucleotides of the guide strand were sequenced using a custom primer (miR30EcoRISeg, TAGCCCCCTTGAATTCGAGGCAGTAGGCA). To provide a sufficient baseline for detecting shRNA depletion in experimental samples, we aimed to acquire >500 reads per shRNA in the T0 sample, which required more than ten million reads per sample to compensate for disparities in shRNA representation inherent in the pooled plasmid preparation or introduced by PCR biases. With these conditions, we acquired T0 baselines of >500 reads for 1,072 (97% of all) shRNAs. Sequence processing was performed using a customized Galaxy platform<sup>32</sup>. For each shRNA and condition, the number of matching reads was normalized to the total number of library-specific reads per lane and imported into a database for further analysis (Access 2003, Microsoft). All shRNA sequences are provided in Supplementary Table 1.

**Animal studies.** All mouse experiments were approved by the Cold Spring Harbor animal care and use committee. Leukaemia cells were transplanted by tail-vein injection of  $1 \times 10^6$  cells into sub-lethally (5.5 Gy) irradiated B6/SJL(CD45.1) recipient mice. For whole-body bioluminescent imaging, mice were intraperitoneally injected with 50 mg kg<sup>-1</sup> D-Luciferin (Goldbio), and after 10 min, analysed using an IVIS Spectrum system (Caliper LifeSciences). Quantification was performed using Living Image software (Caliper LifeSciences) with standardized rectangular regions of interests covering the mouse trunk and extremities. For shRNA induction, animals were treated with doxycycline in both drinking water (2 mg ml<sup>-1</sup> with 2% sucrose; Sigma-Aldrich) and food (625 mg kg<sup>-1</sup>, Harlan laboratories). For JQ1 treatment trials, a stock of 100 mg ml<sup>-1</sup> JQ1 in DMSO was 20-fold diluted by dropwise addition of a 10% 2-hydroxypropyl-β-cyclodextrin carrier (Sigma) under vortexing, yielding a 5 mg ml<sup>-1</sup> final solution. Mice were intraperitoneally injected daily with freshly diluted JQ1 (50 or 100 mg kg<sup>-1</sup>) or a similar volume of carrier containing 5% DMSO.

**Plasmids.** For conditional RNAi experiments, shRNAs were expressed from either the TRMPV-Neo vector or the TtTMPV-Neo vector, which have been described previously (and are available as Addgene catalogue nos 27990 and 27993)<sup>10</sup>. For screen validation, shRNAs were cloned into LMN(MSCV-miR30-PGK-NeoR-IRES-GFP), which was generated from LMP<sup>33</sup> by replacing the PuroR transgene with a NeoR cassette. For Myc rescue experiments, the wild-type mouse Myc cDNA was subcloned into MSCV-PGK-Puro-IRES-GFP (MSCV-PIG)<sup>34</sup>.

**Cell culture.** All mouse MLL-leukaemia cell lines were derived from bone marrow obtained from terminally ill recipient mice, and were cultured in RPMI 1640 (Gibco-Invitrogen) supplemented with 10% FBS, 100 U ml<sup>-1</sup> penicillin and 100 µg ml<sup>-1</sup> streptomycin. MLL-AF9(alone), MLL-AF9/Nras<sup>G12D</sup>, Tet-on MLL-AF9/Nras<sup>G12D</sup> and MLL-ENL/FLT3<sup>ITD</sup> cell cultures were derived as described previously<sup>10,17</sup>. Tet-on immortalized MEF cultures were described previously<sup>10</sup>. G1E cells were provided by M. Weiss. MEF cells were grown in DMEM with 10% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin and 1% glutamine (GIBCO). G1E cells were grown in IMDM with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, 2 U ml<sup>-1</sup> erythropoietin (Sigma) and 10% Kit-ligand-conditioned medium. All human leukaemia cell lines were cultured in RPMI-1640 with 10% FBS, 100 U ml<sup>-1</sup> penicillin and 100 µg ml<sup>-1</sup> streptomycin, except KASUMI-1 cells, which were cultured in 20% FBS. NOMO-1, MOLM-13, EoL-1, NB4, HNT-34 and CMK were purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ). KASUMI-1, HL-60, MV4-11, KG1, HEL, THP-1,

B16-F10 and IMR-90 were obtained from ATCC. K-562 cells were provided by M. Carroll. U2OS, HeLa and Jurkat were provided by the CSHL tissue culture facility. CD34.MA9.NRAS and CD34.MA9.FLT3 cells were generated by retroviral transduction of umbilical-cord-blood CD34<sup>+</sup> cells as described previously<sup>35,36</sup>. All retroviral packaging was performed using ecotropic Phoenix cells according to established protocols ([http://www.stanford.edu/group/nolan/tutorials/retpkg\\_1\\_packlines.html](http://www.stanford.edu/group/nolan/tutorials/retpkg_1_packlines.html)).

**Western blotting.** For Brd4 immunoblotting, 30 µg of whole-cell lysate RIPA extracts (25 mM Tris (pH 7.6), 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) were loaded into each lane. For Myc immunoblotting, cells were lysed directly in Laemmli buffer and about 50,000 cell-equivalents were loaded into each lane. Protein extracts were resolved by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose for blotting.

**Proliferation assay.** Competitive proliferation assays using shRNAs in LMN or TRMPV-/TtTMPV-Neo vectors were performed as outlined in Supplementary Fig. 3a and as described previously<sup>10</sup>, respectively. Proliferation assays for JQ1 *in vitro* testing were performed by counting the increase in viable cell numbers over 72 h in the presence of different JQ1 concentrations. Dead cells were excluded using propidium iodide (PI) staining. Measurements of cell concentration were performed on a Guava EasyCyte (Millipore), gating only viable cells (FSC/SSC/PI<sup>-</sup>). Proliferation rates were calculated using the equation  $\ln(\text{cell concentration at } 72 \text{ h}/\text{cell concentration at } 0 \text{ h})/72$ . Relative proliferation rates were calculated by normalizing to the rate of DMSO-treated cells.

**May-Grünwald-Giemsa cytospin staining.** MLL-AF9/Nras<sup>G12D</sup> leukaemia cells were treated with 1 µg ml<sup>-1</sup> doxycycline for 2 days to induce shRNA expression from TRMPV-Neo or TtTMPV vectors, or treated with 100 nM JQ1 for 2 days. 50,000 cells were resuspended in 100 µl FACS buffer (5% FBS, 0.05% NaN<sub>3</sub> in PBS) and cytospin onto glass slides using a Shandon Cytospin 2 Centrifuge at 500 rpm for 5 min. May-Grünwald (Sigma) and Giemsa (Sigma) stainings were performed according to manufacturer's protocols. Images were collected using a Zeiss Observer Microscope with a ×40 objective.

**BrdU cell-cycle analysis and annexin V flow cytometry.** BrdU incorporation assays were performed according to the manufacturer's protocol (BD, APC BrdU flow kit), with cells pulsed with BrdU for 30 min. Cells were co-stained with 7-aminoactinomycin D or 4',6-diamidino-2-phenylindole (DAPI) for DNA content measurement. For all conditional shRNA experiments, the analysis was gated on Venus<sup>+</sup>/dsRed<sup>+</sup> (shRNA<sup>+</sup>) cell populations. Annexin V apoptosis staining was performed according to the manufacturer's protocol (BD, APC annexin V). To analyse shRNA-mediated induction of apoptosis specifically, annexin V was quantified in viable shRNA-expressing cells (FSC/SSC; Venus<sup>+</sup>/dsRed<sup>+</sup>). Notably, this gating selectively evaluates early apoptotic cells (Annexin V<sup>+</sup>, DAPI<sup>-</sup>), excluding accumulated dead cells (Annexin V<sup>+</sup>, DAPI<sup>+</sup>). All analyses were performed using FlowJo software (Tree Star).

**shRNA experiments in human AML cell lines.** THP-1 and MOLM-13 cells were modified to express the ecotropic receptor and rtTA3 using retroviral transduction of MSCV-RIEP (MSCV-rtTA3-IRES-EcoR-PGK-Puro) followed by drug selection (0.5 and 1 µg ml<sup>-1</sup> puromycin for 1 week, respectively). The resulting cell lines were transduced with ecotropically packaged TRMPV-Neo-shRNA retroviruses, selected with 400 µg ml<sup>-1</sup> G418 for 1 week and treated with 1 µg ml<sup>-1</sup> doxycycline to induce shRNA expression. The relative change in Venus<sup>+</sup>/dsRed<sup>+</sup> (shRNA<sup>+</sup>) cells was monitored on a Guava EasyCyte (Millipore). BrdU cell-cycle analysis was performed as described above.

**Adult primary leukaemia sample analysis.** The study was approved by the Institutional Review Board (ethics committee) of the Medical University of Vienna. Primary leukaemic cells were obtained from peripheral blood or bone marrow aspirate samples. Informed consent was obtained before blood donation or bone marrow puncture in each case. Diagnoses were established according to criteria provided by the French-American-British (FAB) cooperative study group<sup>37,38</sup> and the World Health Organization (WHO)<sup>39</sup>. Mononuclear cells were prepared using Ficoll and stored in liquid nitrogen until used. HL60 and MOLM13 cell lines (obtained from DSMZ) were included as controls. After thawing, the viability of AML cells ranged from 70% to 99% as assessed by trypan blue exclusion. Primary cells (thawed mononuclear cells, 5–10 × 10<sup>4</sup> cells per well) and cell lines (1–5 × 10<sup>4</sup> cells per well) were cultured in 96-well microtitre plates (TPP) in RPMI-1640 medium (PAA laboratories) with 10% fetal calf serum (FCS, Pasching) in the absence or presence of JQ1 (10–5,000 nM) at 37 °C (5% CO<sub>2</sub>) for 48 h. In selected experiments, primary AML cells were incubated with JQ1 in the presence or absence of a cocktail of proliferation-inducing cytokines: recombinant human (rh) G-CSF, 100 ng ml<sup>-1</sup> (Amgen), rhSCF, 100 ng ml<sup>-1</sup> (Peprotech) and rhIL-3, 100 ng ml<sup>-1</sup> (Novartis). After 48 h, 0.5 µCi <sup>3</sup>H-thymidine was added (16 h). Cells were then harvested on filter membranes in a Filtermate 196 harvester (Packard Bioscience). Filters were air-dried and the bound radioactivity was measured in a β-counter (Top-Count NXT, Packard Bioscience). All experiments

were performed in triplicates. Proliferation was calculated as a percentage of control (cells kept in control medium), and the inhibitory effects of JQ1 were expressed as IC<sub>50</sub> values. In 7 out of 12 patients, drug-exposed cells were analysed for morphologic signs of differentiation by Wright–Giemsa staining on cytopspin slides. The thymidine incorporation assay was chosen as a proliferation assay because of its superior sensitivity and ease of implementation for suspension cells, as compared to other proliferation assays, such as MTT.

**Paediatric primary leukaemia sample analysis.** Diagnostic bone marrow samples were collected, under protocols approved by an institutional review board, from newly diagnosed children with acute leukaemia. Informed consent was obtained in accordance with the Helsinki protocol. At the time of collection, primary leukaemic cells were enriched by density centrifugation using Ficoll-Paque PLUS (GE Healthcare) and subsequently stored in liquid nitrogen. Vials of cryopreserved cells were thawed, resuspended in media, and live leukaemic cells were enriched by density centrifugation. Cells were maintained in supplemented media with 20% FBS. All leukaemia cell cultures were incubated at 37 °C in 5% CO<sub>2</sub>. Primary leukaemia samples were treated with dose-ranges of JQ1 and vehicle control for 72 h in 96-well plates. For the annexin binding assays, cells were harvested and stained with Annexin V-PE and 7-AAD (BD Pharmingen), read on a FACSCalibur and analysed with FlowJo software (Tree Star). For the WST-1 assays, WST-1 reagent (Roche Diagnostics) was added to the culture medium (1:10 dilution) and absorbance was measured at 450 nm using a Bio-Rad model 680 microplate reader (Bio-Rad Laboratories). WST-1 assays were performed in triplicate. Primary leukaemia samples were treated with 250 nM JQ1 and vehicle control for 48 h in 96-well plates. Cytopspins were prepared at baseline, 24 h and 48 h and stained with Wright–Giemsa solution (Sigma-Aldrich). Images were acquired using a Nikon Eclipse E600 microscope system (Nikon). Although similar to other metabolic assays measuring cell proliferation (for example, MTT), WST-1 has superior sensitivity for the assessment of cytotoxicity in primary leukaemia samples.

**Histological analysis of bone marrow.** Paraffin-embedded sections were stained with haematoxylin and eosin (H&E). Photographs were taken on a Nikon Eclipse 80i microscope with a Nikon Digital Sight camera using NIS-Elements F2.30 software at a resolution of 2560 × 1920. Using Adobe Photoshop CS2, images were re-sized and set at a resolution of 300 pixels inch<sup>-1</sup>, autocontrast was applied and unsharp mask was used to improve image clarity.

**FACS evaluation of normal haematopoiesis.** Bone marrow cells were obtained by flushing mouse femurs and tibias, followed by erythrocyte lysis with ACK buffer (150 mM NH<sub>4</sub>Cl, 10 mM KHCO<sub>3</sub> and 0.1 mM EDTA). Samples were washed in FACS buffer (5% FBS, 0.05% NaN<sub>3</sub> in PBS), followed by staining (two million cells in 100 µl of FACS buffer) for 1 h. Antibody dilutions used were: mouse haematopoietic lineage eFluor 450 cocktail (1:100), PE-Cy7 anti-mouse Kit (1:50), APC anti-mouse Sca-1 (1:100), APC anti-mouse B220 (1:100), APC anti-mouse Cd11b (1:100), APC anti-mouse TER-119 (1:100) and APC anti-mouse Gr-1 (1:100). Stained samples were analysed on an LSRII flow cytometer. Data analysis was performed using FlowJo software (Treestar).

**Expression microarrays.** Microarrays were performed through the CSHL microarray shared resource. RNA was isolated from 10<sup>7</sup> cells using RNeasy Mini Kit (Qiagen). RNA quality was assessed on an Agilent 2100 Bioanalyser, RNA 6000 Pico Series II Chips (Agilent) and samples with a RIN score of 7.0 or greater passed. RNA was amplified by a modified Eberwine technique, amplified antisense RNA was then converted to cDNA using a WT Expression kit (Ambion). The cDNA was then fragmented and terminally labelled with biotin, using the Affymetrix GeneChip WT Terminal Labelling kit (Affymetrix). Samples were then prepared for hybridization, hybridized, washed and scanned according to the manufacturer's instructions on Mouse Gene ST 1.0 GeneChips (Affymetrix). Affymetrix Expression Console QC metrics were used to pass the image data. Raw data was processed by Affymetrix and Limma packages in R-based Bioconductor. Heatmaps were generated using GenePattern software<sup>40</sup>; RMA-processed microarray data was converted into a log<sub>2</sub> scale, selected gene lists were row-normalized and visualized using the HeatMapImage module on GenePattern. All raw microarray data files are available from the Gene Expression Omnibus (GSE29799).

**GSEA analysis.** Gene set enrichment analysis<sup>20</sup> was performed using GSEA v2.07 software with 1,000 phenotype permutations. Leukaemia-stem-cell and *Myc* gene sets were obtained from the indicated publications<sup>21,22,26</sup>. The macrophage-development gene set was obtained from the Ingenuity Pathway Analysis (IPA) software (Ingenuity). To perform GSEA on human microarray data, mouse gene sets were converted into human gene names using bioDBNet dbWalk (<http://biobdnet.abcc.ncicrf.gov/db/dbWalk.php>) or manually using the NCBI database. A detailed description of GSEA methodology and interpretation is provided at <http://www.broadinstitute.org/gsea/doc/GSEASUserGuideFrame.html>. In brief, the normalized enrichment score (NES) provides 'the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes'. The false discovery rate *q*-value (FDR *q*-val) is 'the estimated probability that a gene set with a given NES represents a false positive finding'. 'In general, given the lack of coherence in most expression datasets and the relatively small number of gene sets being analyzed, an FDR cutoff of 25% is appropriate.' Gene sets used in this study are included in Supplementary Table 2.

**Chromatin immunoprecipitation.** ChIP assays were performed exactly as described<sup>41</sup>. Crosslinking was performed with sequential EGS (Pierce) and formaldehyde<sup>42</sup>. All results were quantified by quantitative PCR performed using SYBR green (ABI) on an ABI 7900HT. Each immunoprecipitate signal was referenced to an input standard-curve dilution series (immunoprecipitate/input) to normalize for differences in starting cell number and for primer amplification efficiency.

**qRT-PCR.** RNA was prepared using TRIzol reagent (Invitrogen). Synthesis of cDNA was performed using qScript cDNA SuperMix (Quanta Biosciences). Quantitative PCR analysis was performed on an ABI 7900HT with SYBR green (ABI). All signals were quantified using the ΔCt method. All signals were normalized to the levels of *Gapdh*.

**Antibodies.** The anti-Brd4 antibody used for western blotting was a gift from G. Blobel. The anti-Brd4 antibody used for ChIP was purchased from Sigma (HPA015055). The anti-Myc antibody was purchased from Epitomics (1472-1). Antibodies used in FACS were: APC anti-mouse CD117/Kit (Biolegend, 105811), APC anti-mouse CD11b (Biolegend, 101211), Pacific Blue anti-mouse CD45.2 (Biolegend, 109820), mouse haematopoietic lineage eFluor 450 cocktail (eBioscience, 88-7772-72), APC anti-mouse CD45R/B220 (Biolegend, 103212), APC anti-mouse TER-119/erythroid cells (Biolegend, 116212), APC anti-mouse Ly-6G/Gr-1 (eBioscience, 17-5931), PE-Cy7 anti-mouse CD117/Kit (eBioscience, 25-1171-82) and APC anti-mouse Sca-1 (eBioscience, 17-5981-81). The anti-β-actin HRP antibody was purchased from Sigma (A3854).

32. Taylor, J., Schenck, I., Blankenberg, D. & Nekrutenko, A. Using galaxy to perform large-scale interactive data analyses in *Curr. Protoc. Bioinformatics* Ch. 10, Unit 10.5 (2007).
33. Dickens, R. A. *et al.* Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nature Genet.* **37**, 1289–1295 (2005).
34. Hemann, M. T. *et al.* An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nature Genet.* **33**, 396–400 (2003).
35. Wei, J. *et al.* Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. *Cancer Cell* **13**, 483–495 (2008).
36. Wunderlich, M. & Mulloy, J. C. Model systems for examining effects of leukemia-associated oncogenes in primary human CD34<sup>+</sup> cells via retroviral transduction. *Methods Mol. Biol.* **538**, 263–285 (2009).
37. Bennett, J. M. *et al.* Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br. J. Haematol.* **33**, 451–458 (1976).
38. Bennett, J. M. *et al.* Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British cooperative group. *Ann. Intern. Med.* **103**, 620–625 (1985).
39. Vardiman, J. W. *et al.* The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* **114**, 937–951 (2009).
40. Reich, M. *et al.* GenePattern 2.0. *Nature Genet.* **38**, 500–501 (2006).
41. Steger, D. J. *et al.* DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol. Cell. Biol.* **28**, 2825–2839 (2008).
42. Zeng, P. Y., Vakoc, C. R., Chen, Z. C., Blobel, G. A. & Berger, S. L. *In vivo* dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. *Biotechniques* **41**, 694–698 (2006).

# Inhibition of BET recruitment to chromatin as an effective treatment for MLL–fusion leukaemia

Mark A. Dawson<sup>1,2\*</sup>, Rab K. Prinjha<sup>3\*</sup>, Antje Dittmann<sup>4\*</sup>, George Giotopoulos<sup>1</sup>, Marcus Bantscheff<sup>4</sup>, Wai-In Chan<sup>1</sup>, Samuel C. Robson<sup>2</sup>, Chun-wa Chung<sup>5</sup>, Carsten Hopf<sup>4</sup>, Mikhail M. Savitski<sup>4</sup>, Carola Huthmacher<sup>4</sup>, Emma Gudgin<sup>1</sup>, Dave Lugo<sup>3</sup>, Soren Beinke<sup>3</sup>, Trevor D. Chapman<sup>3</sup>, Emma J. Roberts<sup>3</sup>, Peter E. Soden<sup>3</sup>, Kurt R. Auger<sup>6</sup>, Olivier Mirguet<sup>7</sup>, Konstanze Doehner<sup>3</sup>, Ruud Delwel<sup>9</sup>, Alan K. Burnett<sup>10</sup>, Phillip Jeffrey<sup>3</sup>, Gerard Drewes<sup>4</sup>, Kevin Lee<sup>3</sup>, Brian J. P. Huntly<sup>1\*</sup> & Tony Kouzarides<sup>2\*</sup>

Recurrent chromosomal translocations involving the mixed lineage leukaemia (MLL) gene initiate aggressive forms of leukaemia, which are often refractory to conventional therapies<sup>1</sup>. Many MLL-fusion partners are members of the super elongation complex (SEC), a critical regulator of transcriptional elongation, suggesting that aberrant control of this process has an important role in leukaemia induction<sup>2,3</sup>. Here we use a global proteomic strategy to demonstrate that MLL fusions, as part of SEC<sup>2,3</sup> and the polymerase-associated factor complex (PAF)<sup>4,5</sup>, are associated with the BET family of acetyl-lysine recognizing, chromatin ‘adaptor’ proteins. These data provided the basis for therapeutic intervention in MLL-fusion leukaemia, via the displacement of the BET family of proteins from chromatin. We show that a novel small molecule inhibitor of the BET family, GSK1210151A (I-BET151), has profound efficacy against human and murine MLL-fusion leukaemic cell lines, through the induction of early cell cycle arrest and apoptosis. I-BET151 treatment in two human leukaemia cell lines with different MLL fusions alters the expression of a common set of genes whose function may account for these phenotypic changes. The mode of action of I-BET151 is, at least in part, due to the inhibition of transcription at key genes (*BCCL2*, *C-MYC* and *CDK6*) through the displacement of BRD3/4, PAFc and SEC components from chromatin. *In vivo* studies indicate that I-BET151 has significant therapeutic value, providing survival benefit in two distinct mouse models of murine MLL–AF9 and human MLL–AF4 leukaemia. Finally, the efficacy of I-BET151 against human leukaemia stem cells is demonstrated, providing further evidence of its potent therapeutic potential. These findings establish the displacement of BET proteins from chromatin as a promising epigenetic therapy for these aggressive leukaemias.

Dysregulation of chromatin modifiers is a recurrent and sentinel event in oncogenesis<sup>6</sup>. Therapeutic strategies that selectively alter the recruitment and/or catalytic activity of these enzymes at chromatin therefore hold great promise as targeted therapies<sup>6</sup>. In this regard the bromodomain and extra terminal (BET) family of proteins (BRD2, BRD3, BRD4 and BRDT) provide an ideal ‘druggable’ target, because they share a common highly conserved tandem bromodomain at their amino terminus. Selective bromodomain inhibitors that disrupt the binding of BET proteins to histones have recently been described<sup>7,8</sup>; however, their true therapeutic scope remains untested.

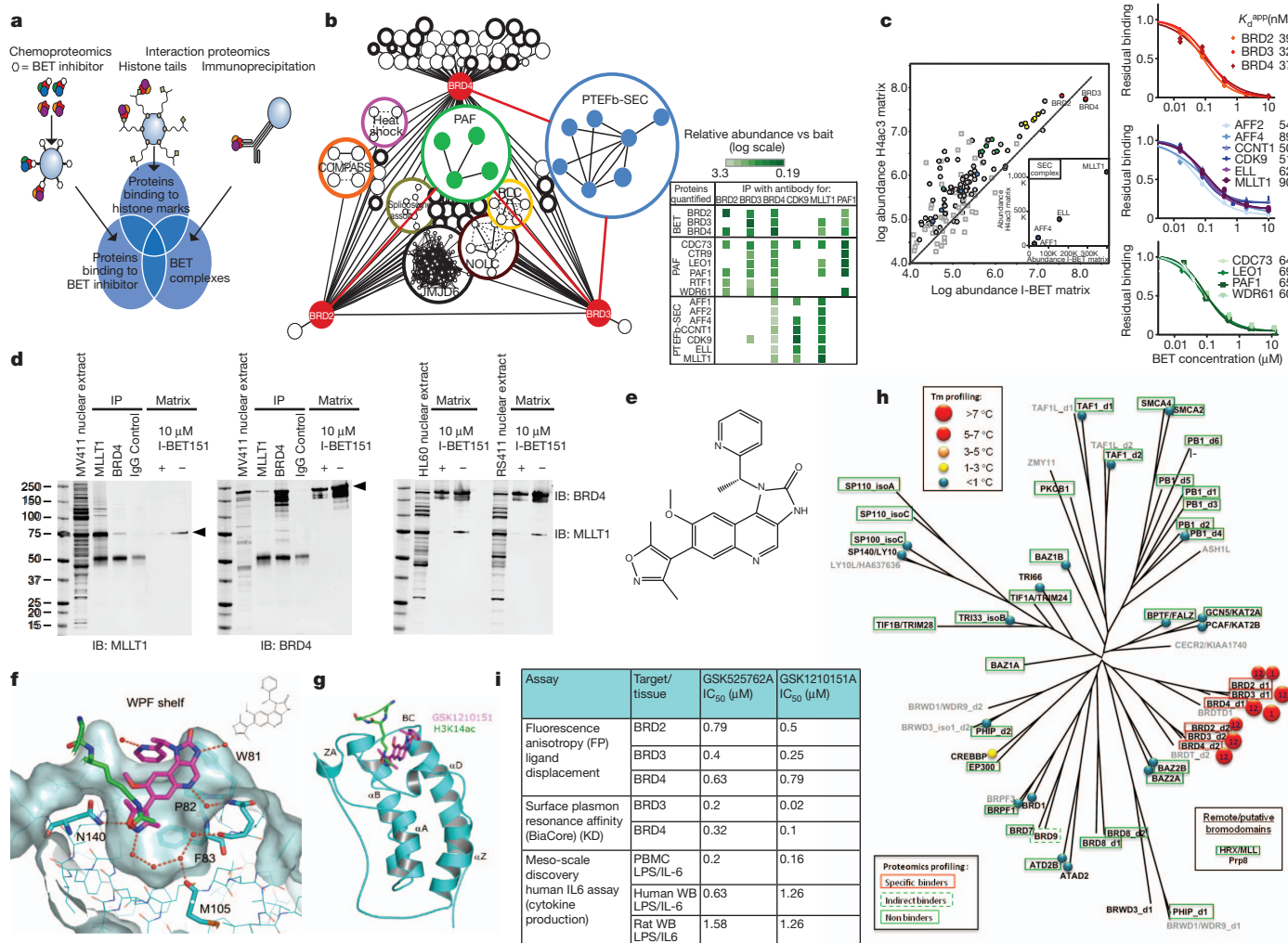
To identify the nuclear complexes associated with ubiquitously expressed BETs (BRD2/3/4), we performed a systematic global proteomic survey. Specifically, this involved a tri-partite discovery approach (Fig. 1a). In the first approach, bead-immobilized analogues

of I-BET762 (ref. 9) were incubated with HL60 nuclear extracts and bound proteins were analysed by quantitative mass spectrometry (Supplementary Table 1). This approach identified the BET isoforms and a large number of co-purifying proteins (Supplementary Tables 1 and 2), indicating that the BET isoforms reside in many distinct protein complexes. In the second approach, immunoprecipitation analyses with selective antibodies against BRD2/3/4 were performed (Supplementary Fig. 1 and Supplementary Tables 3 and 4). This was complemented with additional immunoprecipitations using selected antibodies against complex members (‘baits’) selected from the subset of proteins that were identified in the first approach (Fig. 1b right panel, Supplementary Fig. 2 and Supplementary Table 3). In the third approach, bead-immobilized histone H4(1–21; K5acK8acK12ac) acetylated peptides were used to purify protein complexes. These data were combined to highlight a list of complexes identified in all three methods (Fig. 1b left panel, Supplementary Fig. 3 and Supplementary Table 1). Finally, specificity of the I-BET762 and histone tail matrix was further assessed by competition experiments (Fig. 1c, Supplementary Figs 4, 5 and Supplementary Table 2). This strategy enabled the direct determination of the targets of the inhibitor, and the proteins associated with the target, with subunits of protein complexes exhibiting closely matching half-maximum inhibitory concentration (IC<sub>50</sub>) values<sup>10</sup>. Taken together these stringent and complementary approaches provide a high confidence global data set encompassing all known<sup>11–13</sup> and several novel BET protein complexes (Fig. 1b and Supplementary Fig. 3). Among the novel complexes, we observed a prominent enrichment and dose-dependent inhibition of several components of the PAFc<sup>4,5</sup> and SEC<sup>2,3</sup> (Fig. 1b, c), which were confirmed by reciprocal immunoprecipitations in HL60 cells (Fig. 1b). Moreover, reciprocal immunoprecipitations in two MLL-fusion leukaemia cell lines (MV4;11 and RS4;11) confirmed the relationship of SEC with BRD4 in different cellular contexts (Fig. 1d). Together these data indicate that BRD3 and BRD4 associate with the PAFc and SEC and may function to recruit these complexes to chromatin. Given that these complexes are crucial for malignant transformation by MLL fusions<sup>2–5</sup> we tested the hypothesis that displacement of BET proteins from chromatin may have a therapeutic role in these leukaemias.

To progress our studies with an optimized therapeutic agent we developed I-BET151 (Fig. 1e); a novel dimethylisoxazole template, previously undisclosed as a BET bromodomain inhibitor. It was identified and optimized to retain excellent BET target potency (Fig. 1i) and selectivity (Fig. 1h, Supplementary Figs 5–10 and Supplementary Table 5) while enhancing the *in vivo* pharmacokinetics and terminal half-life to enable prolonged *in vivo* studies (Fig. 4a and Supplementary

<sup>1</sup>Department of Haematology, Cambridge Institute for Medical Research and Addenbrookes Hospital, University of Cambridge, Cambridge CB2 0XY, UK. <sup>2</sup>Gurdon Institute and Department of Pathology, Tennis Court Road, Cambridge CB2 1QN, UK. <sup>3</sup>Epinova DPU, Immuno-Inflammation Centre of Excellence for Drug Discovery, GlaxoSmithKline, Medicines Research Centre, Gunnels Wood Road, Stevenage SG1 2NY, UK. <sup>4</sup>Cellzome AG Meyerhofstrasse 1, 69117 Heidelberg, Germany. <sup>5</sup>Molecular Discovery Research, GlaxoSmithKline R&D, Stevenage SG1 2NY, UK. <sup>6</sup>Cancer Epigenetics DPU, Oncology R&D, GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, Pennsylvania 19426, USA. <sup>7</sup>Lipid Metabolism Discovery Performance Unit, GSK R&D, 91951 Les Ulis Cedex, France. <sup>8</sup>University Hospital of Ulm Internal Medicine III Albert-Einstein-Allee 23, 89081 Ulm, Germany. <sup>9</sup>Department of Hematology, Erasmus University Medical Center, 3015 GE Rotterdam, The Netherlands. <sup>10</sup>Department of Hematology, Cardiff University School of Medicine, Cardiff, CF14 4XN, UK.

\*These authors contributed equally to this work.



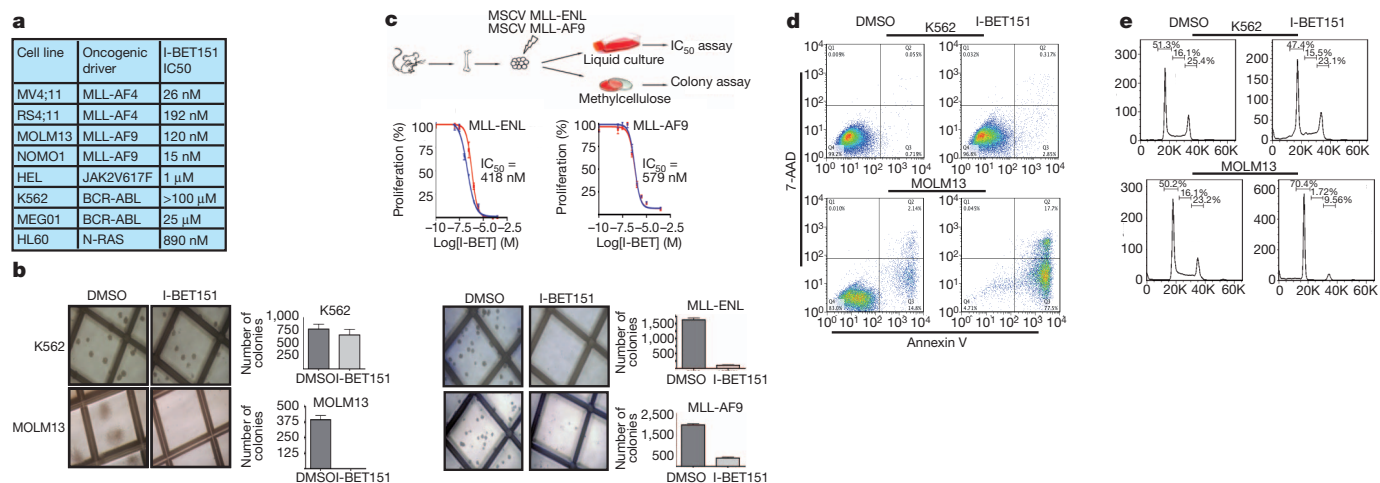
**Figure 1 | A global proteomic survey identifies BET proteins as part of the PAFc and SEC.** **a**, Proteomic strategy. **b**, Left, Cytoscape representation of the BET protein complex network (discussed in detail in Supplementary Fig. 3). Bold circles indicate associations confirmed by the three orthogonal methods. Right, Heat map representing quantitative-mass spectrometry data following co-immunoprecipitation of BETs, PAF and SEC complex members. **c**, Differential proteomic analysis of the proteins interacting with I-BET and triple acetylated histone H4 tail. Left, affinity matrices with immobilized I-BET762 or histone H4(K5acK8acK12ac) peptide bind to the same set of BET complexes. Protein abundance was determined from signal intensities in the mass spectrometer (arbitrary units,  $K = \times 1,000$ ). Right, competitive inhibition of the binding of BET isoforms, and SEC and PAF complex components, to the I-BET762 matrix showing matching concentration dependence. **d**, BRD4 and MLLT1 interact in HL60, MV4;11 and RS4;11 cells and binding to the I-BET762 matrix is blocked by excess I-BET151. **e**, Chemical structure of GSK1210151A (I-BET151). **f**, I-BET151 binding to the acetyl-binding pocket of BRD4-BD1 (cyan) overlaid with H3K14-acetyl peptide (green) (Protein

Database ID 3jvk). A surface representation of the BRD4-BD1 is shown with key recognition and the specificity WPF shelf identified. **g**, Ribbon representation of the BRD4-BD1 (cyan) crystal structure complexed with I-BET151 (shown in magenta stick format) overlaid with H3(12-19)K14ac peptide (green) taken from its complex with BRD4-BD1 (PDB ID 3jvk). Secondary elements of the BRD4-BD1 structure have been highlighted. **h**, Selectivity profile of I-BET-151 showing average temperature shifts ( $T_m$ ) using a fluorescent thermal shift assay. Numbering inside the spheres indicates bromodomains assessed; for example, 12 signifies both bromodomains 1 and 2 have been assessed. Overlaid is the selectivity profile generated using a proteomic approach (shown as boxes around proteins, discussed in Supplementary Fig. 5). Where the bromodomains have been profiled by both thermal shift and proteomic approaches the agreement is excellent. Proteins not assessed by either technique are shown in grey. **i**, Comparison of I-BET762 and I-BET151 potency in ligand displacement assays, direct Biacore binding and lipopolysaccharide-stimulated IL-6 cytokine production from human peripheral blood mononuclear cells (PBMC) or whole blood (WB).

Fig. 20). We also generated proteomic selectivity profiles comparing I-BET151 with I-BET762 (Fig. 1h, Supplementary Fig. 5 and Supplementary Table 6). We bead-immobilized a combination of differentially acetylated histone tail peptides (Supplementary Table 7), which captured a total of 27 bromodomain proteins from HL60 nuclear extracts. Competition with excess I-BET151 or I-BET762 blocked the capture of BRD2, BRD3, BRD4, and BRD9 but had no effect on the 23 other bromodomain proteins including MLL. The inhibition of BRD9 is likely to be indirect as this protein forms a complex with BRD4 (Supplementary Table 3). Finally, a high-resolution (1.5 Å) crystal structure of I-BET151 bound to BRD4-bromodomain 1

(BD1) revealed binding to the acetylated-lysine (AcK) recognition pocket of the BET protein (Fig. 1f, g and Supplementary Fig. 10).

To assess the therapeutic efficacy and selectivity of I-BET151, we tested a panel of leukaemic cell lines harbouring a spectrum of distinct oncogenic drivers. These data demonstrated that I-BET151 has potent efficacy against cell lines harbouring different MLL-fusions (Fig. 2a and Supplementary Fig. 11). To extend these data we tested the clonogenic potential of human leukaemic cells grown in cytokine-supplemented methylcellulose containing dimethylsulphoxide (DMSO; vehicle) or I-BET151. Consistent with the profound effects in liquid culture, the colony-forming potential of MLL-fusion-driven



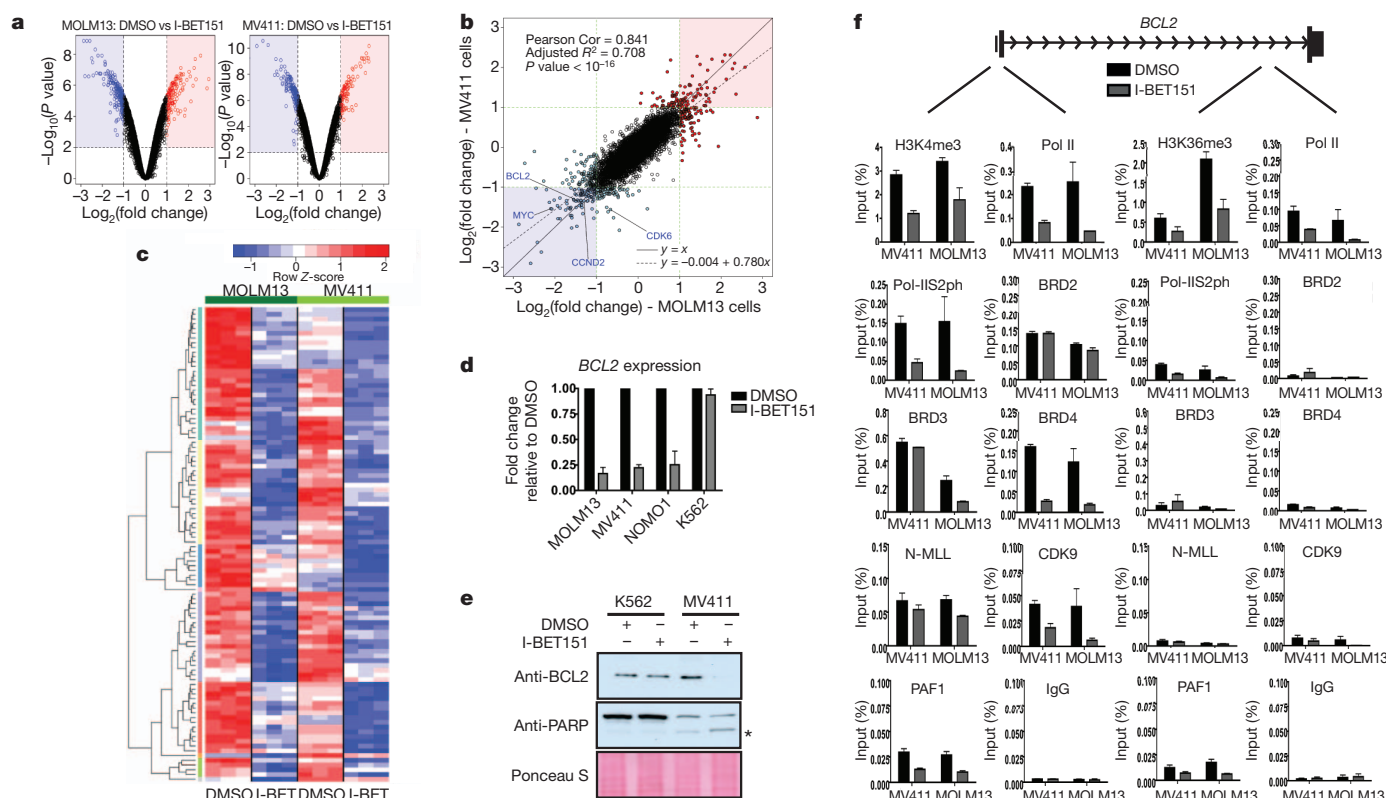
**Figure 2 | I-BET151 selectively and potently inhibits MLL-fusion leukaemic cell lines *in vitro*.** **a**, Human leukaemia cell lines tested using I-BET151. **b**, Clonogenic assays performed in the presence of DMSO or I-BET151. **c**, Haematopoietic progenitors were isolated from mouse bone marrow and retrovirally transformed with MLL-ENL or MLL-AF9. These cells were used in both proliferation and clonogenic assays **d**, Apoptosis was assessed by FACS

analysis after 72 h incubation with DMSO or I-BET151. **e**, Cell cycle progression was assessed by FACS analysis 24 h after incubation with DMSO or I-BET151 (y axis event count, x axis arbitrary fluorescence units). Bar graphs are represented as the mean and error bars reflect standard deviation of results derived from triplicate experiments.

leukaemias (MOLM13) was completely ablated by I-BET151, whereas leukaemias driven by tyrosine kinase activation (K562) were unaffected (Fig. 2b). In addition to the data with human leukaemic cell lines, we also confirmed the potent efficacy of I-BET151 in both liquid culture and clonogenic assays using primary murine progenitors

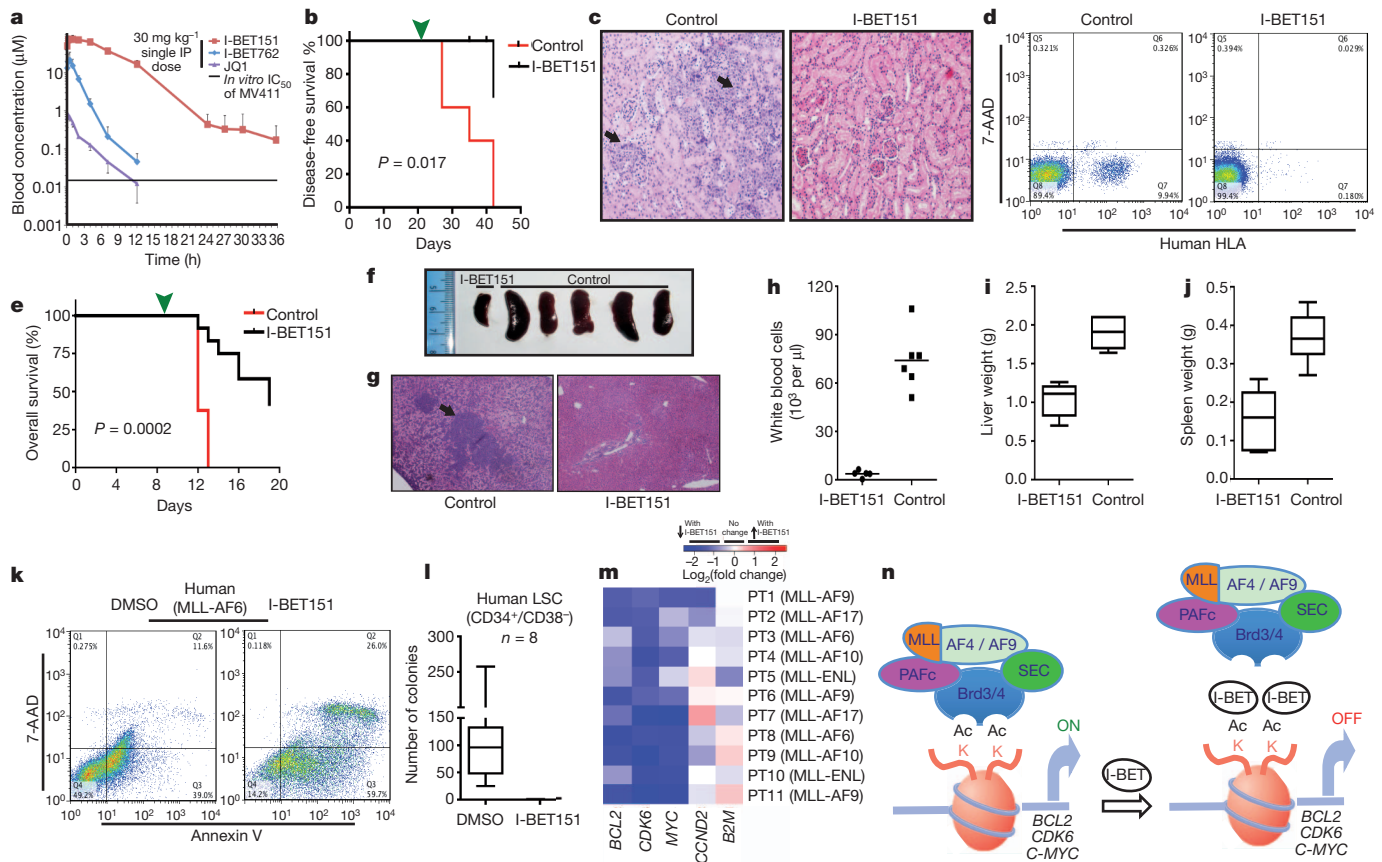
retrovirally transformed with either MLL-ENL or MLL-AF9 (Fig. 2c).

To investigate the mechanism of action for I-BET151, we performed fluorescence-activated cell sorting (FACS) analysis to assess apoptosis and cell cycle progression after I-BET151 treatment. Figure 2d–e and Supplementary Fig. 12 show a marked induction of apoptosis and a



**Figure 3 | Transcriptome and ChIP analyses provide mechanistic insights for the efficacy of I-BET151.** **a**, Volcano plots for DMSO against I-BET151 treated samples, showing the adjusted significance  $P$  value ( $\log_{10}$ ) versus fold change ( $\log_2$ ). **b**, Correlation of  $\log_2$  fold change between MV411 and MOLM13 across all genes. No genes show opposing expression changes. Lines represent the identity line (black solid), the line of best fit (black dotted), or  $\log_2$  fold-change threshold values (green dotted). **c**, Heat map of top 100 genes

downregulated following treatment with I-BET151. **d**,  $BCL2$  gene expression (normalized to  $B2M$  expression) is shown. Expression level of  $BCL2$  in DMSO was assigned a value of 1. **e**, Immunoblotting demonstrating a decrease in  $BCL2$  and an increase in cleaved PARP (\*) after I-BET151 treatment. **f**, ChIP analysis at the TSS and 3' end of  $BCL2$  is illustrated. Bar graphs are represented as the mean enrichment relative to input and error bars reflect standard deviation of results derived from biological triplicate experiments.



**Figure 4 | I-BET151 is efficacious in *in vivo* murine models and primary patient samples of MLL-fusion leukaemia.** **a**, Murine pharmacokinetic studies (mean  $\pm$  s.d.,  $n = 4$  per compound) comparing the blood concentration of I-BET151 with I-BET762 and JQ1. **b**, Kaplan-Meier curve of control and treated NOD-SCID mice transplanted with  $1 \times 10^7$  MV4;11 cells. Green arrowhead, treatment commencement on day 21. **c**, Haematoxylin and eosin-stained histological sections of the renal parenchyma of control and treated mice. Black arrows highlight leukaemic infiltration. **d**, Representative FACS analysis from the peripheral blood of control or I-BET151-treated mice. **e**, Kaplan-Meier curve of control and treated C57BL/6 mice transplanted with  $2.5 \times 10^6$  syngeneic MLL-AF9 leukaemic cells. Green arrowhead, treatment commencement on day 9. **f**, Photomicrograph of the spleen size from 5/8 control and 1/12 I-BET151-treated mice that died on day 12. **g**, Haematoxylin

and eosin-stained histological sections of the liver parenchyma from control and I-BET151-treated mice demonstrating reduced disease burden in the treated animal. **h-j**, Peripheral blood white cell count (**h**), liver weight (**i**) and spleen weights (**j**) from all the control and treated mice at the time of necropsy. **k**, Representative FACS analysis assessing apoptosis from a patient with MLL-AF6 leukaemia. **l**, Clonogenic assays with human MLL-fusion LSC isolated by FACS sorting ( $\text{CD}34^+/\text{CD}38^-$ ) and plated in the presence of DMSO or I-BET151. **m**, Gene expression changes in human MLL-fusion leukaemia cells following treatment with I-BET151 or DMSO. The log<sub>2</sub> fold change in the expression level for all genes (expression level with I-BET151 treatment/expression level with DMSO) is represented. **n**, Schematic model proposing the mode of action for I-BET151 in MLL-fusion leukaemia.

prominent G<sub>0</sub>/G<sub>1</sub> arrest in two MLL-fusion cell lines driven by distinct MLL fusions (MOLM13 and MV4;11 containing MLL-AF9 and MLL-AF4, respectively). In contrast, the cell cycle characteristics and apoptotic rate of K562 cells were largely unaffected at this time. These data indicate that I-BET151 alters the transcriptional programmes regulating apoptosis and cell-cycle progression in MLL-fusion leukaemias. To identify the precise transcriptional pathways controlled by I-BET151, global gene-expression analysis was performed in MOLM13 and MV4;11 cells after treatment with I-BET151 or DMSO for 6 h. This strategy allowed us to identify early I-BET151-responsive genes, before any discernable phenotypic alteration in cell cycle or apoptosis (Supplementary Fig. 12). As demonstrated previously<sup>7</sup>, we observed differential expression of a selective subset of genes (Fig. 3a), rather than global transcriptional dysregulation. Remarkably, the transcriptional programmes altered in the two MLL-fusion cell lines were highly correlated (Fig. 3b) and gene set enrichment analysis documented significant overlap with published MLL fusion signatures including MLL-fusion leukaemia stem cells (LSC)<sup>14,15</sup> (Supplementary Fig. 13). These data are consistent with the notion that MLL fusions aberrantly co-opt the SEC and PAFc to regulate similar transcriptional

programmes. Notably, the top 100 genes concomitantly decreased in both MOLM13 and MV4;11 (Fig. 3c) contained several previously reported direct MLL targets, such as *BCL2*, *CDK6* and *MYC*, the down-regulation of which was consistent with the phenotypic consequences of I-BET151 treatment. *BCL2* is a key antiapoptotic gene implicated in the pathogenesis of MLL-fusion leukaemias<sup>16,17</sup>. Consistent with these data, I-BET151 reduced the expression of *BCL2* in a third MLL-fusion cell line (NOMO1) but not in the unresponsive K562 cells (Fig. 3d), and induction of apoptosis coincided with a marked reduction in *BCL2* protein expression (Fig. 3e). Moreover, overexpression of *BCL2* in the presence of I-BET151 rescued the apoptotic phenotype (Supplementary Fig. 14). Chromatin immunoprecipitation (ChIP) analyses at the *BCL2* locus showed that 6 h of I-BET151 treatment selectively decreased the recruitment of BRD3/4 and impaired recruitment of CDK9 and PAF1 (part of SEC and PAFc, respectively) to the transcriptional start site (TSS). This correlated with reduced phosphorylation of RNA polymerase II (Pol II) on serine 2 of its carboxy-terminal domain (Pol-II-S2ph) (Fig. 3f). A similar pattern was observed at two other MLL target genes (*MYC* and *CDK6*), but not at housekeeping genes (*B2M*) whose expression was unaltered by I-BET151 (Supplementary Fig. 15).

programmes. Notably, the top 100 genes concomitantly decreased in both MOLM13 and MV4;11 (Fig. 3c) contained several previously reported direct MLL targets, such as *BCL2*, *CDK6* and *MYC*, the down-regulation of which was consistent with the phenotypic consequences of I-BET151 treatment.

*BCL2* is a key antiapoptotic gene implicated in the pathogenesis of MLL-fusion leukaemias<sup>16,17</sup>. Consistent with these data, I-BET151 reduced the expression of *BCL2* in a third MLL-fusion cell line (NOMO1) but not in the unresponsive K562 cells (Fig. 3d), and induction of apoptosis coincided with a marked reduction in *BCL2* protein expression (Fig. 3e). Moreover, overexpression of *BCL2* in the presence of I-BET151 rescued the apoptotic phenotype (Supplementary Fig. 14). Chromatin immunoprecipitation (ChIP) analyses at the *BCL2* locus showed that 6 h of I-BET151 treatment selectively decreased the recruitment of BRD3/4 and impaired recruitment of CDK9 and PAF1 (part of SEC and PAFc, respectively) to the transcriptional start site (TSS). This correlated with reduced phosphorylation of RNA polymerase II (Pol II) on serine 2 of its carboxy-terminal domain (Pol-II-S2ph) (Fig. 3f). A similar pattern was observed at two other MLL target genes (*MYC* and *CDK6*), but not at housekeeping genes (*B2M*) whose expression was unaltered by I-BET151 (Supplementary Fig. 15).

Together, these data indicate that the mechanism of efficacy for I-BET151 involves a selective abrogation of BRD3/4 recruitment to chromatin. The consequence of this is the inefficient phosphorylation/recruitment of Pol II. Further investigation is necessary to distinguish whether Pol II recruitment and/or elongation is primarily affected by I-BET151.

We next sought to establish the therapeutic potential of I-BET151 *in vivo*. We first characterized the pharmacokinetic properties of I-BET151 in several preclinical species (Supplementary Fig. 20) and also compared it to published inhibitors<sup>7,8</sup> (Fig. 4a). We then assessed the efficacy of I-BET151 in two established models of MLL leukaemia. Our first model was a xenotransplant model of disseminated human MLL–AF4 leukaemia<sup>18</sup>. I-BET151 was delivered daily at 30 mg kg<sup>−1</sup> by intraperitoneal injection from day 21 (ref. 18), and mice were humanely killed if clinical disease dictated or if there was a sequential rise in peripheral blood disease. At the experimental end-point all the control mice had succumbed to fulminant or progressive disease whereas only one out of five mice in the treated cohort had evidence of disease at low levels (Fig. 4b–d and Supplementary Fig. 16). In our second syngeneic model of murine MLL–AF9 leukaemia, 2.5 × 10<sup>6</sup> leukaemic cells, established from serial transplantation, were injected into tertiary recipients. Despite the latency being reduced to less than 15 days, we waited to initiate treatment from day 9 to test the efficacy of I-BET151 in the setting of overwhelming established disease (Fig. 4e), the scenario often encountered in clinical practice. Even here I-BET151 provided a clear and marked survival benefit (Fig. 4e–j and Supplementary Fig. 17). Taken together, these data demonstrate that I-BET151 provides excellent control of MLL leukaemia progression in two distinct and complementary murine models.

Finally, to demonstrate the applicability of our findings to human disease, we tested the efficacy of I-BET151 in leukaemia cells isolated from patients with various MLL fusions. These data show that I-BET151 accelerates apoptosis (Fig. 4k and Supplementary Fig. 18), and abrogates clonogenic efficiency in bulk leukaemia (Supplementary Fig. 19) as well as isolated LSC (Fig. 4l). These effects are driven, at least in part, by downregulation of a similar transcription programme identified in MLL-fusion cell lines (Fig. 4m). Taken together, these data provide compelling evidence of therapeutic potential and suggest that disease eradication is possible.

The paradigm for epigenetic drug discovery shown here highlights an emerging role for targeting aberrant transcriptional elongation in oncogenesis<sup>2–5</sup> and provides the first example in epigenetic therapy where mechanistic insights have driven targeted drug discovery and application (Fig. 4n). Together, our results suggest that perturbing the interaction of BET proteins with chromatin using I-BET151 may be of great therapeutic value in human MLL-fusion leukaemias. Using a complementary strategy and a different BET inhibitor, a separate study published in this issue concurs with this view<sup>19</sup>. Moreover, the extensive proteomic resource provided here has identified other important disease-associated proteins binding to BET proteins, such as MMSET (WHSC1), which is implicated in multiple myeloma<sup>20</sup>. This raises the possibility that BET inhibitors may have an even wider therapeutic scope in oncology and perhaps in other areas of unmet need within the clinical arena.

## METHODS SUMMARY

Cell culture, gene expression, chromatin immunoprecipitation and FACS analysis were performed as previously described<sup>21</sup>. Proteomic profiling and characterization of inhibitor specificity was performed using methodology previously described<sup>7,9,10</sup>. Detailed information about the reagents and methodology used in this study is available in Supplementary Information.

Received 29 June; accepted 30 August 2011.

Published online 2 October; corrected 27 October 2011 (see full-text HTML version for details).

1. Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nature Rev. Cancer* **7**, 823–833 (2007).

2. Lin, C. *et al.* AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell* **37**, 429–437 (2010).
3. Yokoyama, A., Lin, M., Nares, A., Kitabayashi, I. & Cleary, M. L. A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell* **17**, 198–212 (2010).
4. Milne, T. A. *et al.* Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol. Cell* **38**, 853–863 (2010).
5. Muntean, A. G. *et al.* The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell* **17**, 609–621 (2010).
6. Rodríguez-Paredes, M. & Esteller, M. Cancer epigenetics reaches mainstream oncology. *Nature Med.* **17**, 330–339 (2011).
7. Nicodeme, E. *et al.* Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119–1123 (2010).
8. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
9. Chung, C. W. *et al.* Discovery and characterization of small molecule inhibitors of the BET family bromodomains. *J. Med. Chem.* **54**, 3827–3838 (2011).
10. Bantscheff, M. *et al.* Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nature Biotechnol.* **29**, 255–265 (2011).
11. Jang, M. K. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell* **19**, 523–534 (2005).
12. Maruyama, T. *et al.* A mammalian bromodomain protein, Brd4, interacts with replication factor C and inhibits progression to S phase. *Mol. Cell. Biol.* **22**, 6509–6520 (2002).
13. Yang, Z. *et al.* Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell* **19**, 535–545 (2005).
14. Somerville, T. C. *et al.* Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. *Cell Stem Cell* **4**, 129–140 (2009).
15. Wang, J. *et al.* Conditional MLL-CBP targets GMP and models therapy-related myeloproliferative disease. *EMBO J.* **24**, 368–381 (2005).
16. Robinson, B. W. *et al.* Abundant anti-apoptotic BCL-2 is a molecular target in leukaemias with t(4;11) translocation. *Br. J. Haematol.* **141**, 827–839 (2008).
17. Wang, Q. F. *et al.* MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood* **117**, 6895–6905 (2011).
18. O'Farrell, A. M. *et al.* SU11248 is a novel FLT3 tyrosine kinase inhibitor with potent activity *in vitro* and *in vivo*. *Blood* **101**, 3597–3605 (2003).
19. Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* doi:10.1038/nature10334 (this issue).
20. Martínez-García, E. *et al.* The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood* **117**, 211–220 (2011).
21. Dawson, M. A. *et al.* JAK2 phosphorylates histone H3Y41 and excludes HP1α from chromatin. *Nature* **461**, 819–822 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank S. J. Dawson, A. Bannister, S. Anand and all members of the Huntly and Kouzarides laboratories. We are grateful to H. Doehner, the NCRI AML trials biobank and A. Giles for the provision of patient samples. We acknowledge D. Huang for the BCL2 expression plasmid, L. Gordon for supplying fluorescence resonance energy transfer data and R. Woodward, C. Delves, E. Jones and P. Holmes for protein production. J. Witherington, N. Parr, S. Baddeley and J. Seal provided compound selectivity data. We thank N. Deeks and L. Cutler for providing sample and PK data analysis. We acknowledge K. Smitheman and A. Wyce for help with the cellular analysis of the BET inhibitors, P. Grandi for suggestions and discussion, S. Chan for biophysical assay data, and members of the Epinova team for discussion and suggestions. We thank staff at the ESRF at Grenoble for beamline assistance. We thank T. Werner for assistance with mass spectrometry experiments and data analysis, and the members of the Cellzome Biochemistry, Mass Spectrometry, and IT teams for outstanding expertise and diligence. This work was supported by a Wellcome-Beit Intermediate Clinical Fellowship to M.A.D. The Huntly lab is funded by the Medical Research Council (UK), Leukaemia Lymphoma Research (UK), the Wellcome Trust, The Leukemia & Lymphoma Society of America, Cancer Research UK (CRUK) and the NIHR Cambridge Biomedical Research Centre. This work in the Kouzarides laboratory was funded by a programme grant from Cancer Research UK (CRUK).

**Author Contributions** M.A.D., R.K.P., A.D., G.D., K.L., P.J., B.J.P.H. and T.K. designed the research, interpreted data and wrote the manuscript. M.A.D., A.D., G.G., M.B., W.-I.C., S.C.R., C.-w.C., C.H., M.M.S., C.H., E.G., D.L., S.B., T.D.C., E.J.R., P.E.S., K.R.A. and O.M. performed experiments and analysed data. K.D., R.D. and A.K.B. provided patient samples. M.A.D., R.K.P. and A.D. are joint first authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). I-BET151 compound requests should be directed to K.L. (kevin.2.lee@gsk.com). Correspondence and requests for materials should be addressed to B.J.P.H. (bjph2@cam.ac.uk), M.A.D. (maf2@cam.ac.uk) or T.K. (t.kouzarides@gurdon.cam.ac.uk).

# Active-site remodelling in the bifunctional fructose-1,6-bisphosphate aldolase/phosphatase

Juan Du<sup>1\*</sup>, Rafael F. Say<sup>2\*</sup>, Wei Lü<sup>1</sup>, Georg Fuchs<sup>2</sup> & Oliver Einsle<sup>1,3</sup>

**Fructose-1,6-bisphosphate (FBP) aldolase/phosphatase is a bifunctional, thermostable enzyme that catalyses two subsequent steps in gluconeogenesis in most archaea and in deeply branching bacterial lineages<sup>1–3</sup>. It mediates the aldol condensation of heat-labile dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (GAP) to FBP<sup>4</sup>, as well as the subsequent, irreversible hydrolysis of the product to yield the stable fructose-6-phosphate (F6P) and inorganic phosphate; no reaction intermediates are released. Here we present a series of structural snapshots of the reaction that reveal a substantial remodelling of the active site through the movement of loop regions that create different catalytic functionalities at the same location. We have solved the three-dimensional structures of FBP aldolase/phosphatase from thermophilic *Thermoproteus neutrophilus*<sup>5,6</sup> in a ligand-free state as well as in complex with the substrates DHAP and FBP and the product F6P to resolutions up to 1.3 Å. In conjunction with mutagenesis data, this pinpoints the residues required for the two reaction steps and shows that the sequential binding of additional Mg<sup>2+</sup> cations reversibly facilitates the reaction. FBP aldolase/phosphatase is an ancestral gluconeogenic enzyme optimized for high ambient temperatures<sup>1,2</sup>, and our work resolves how consecutive structural rearrangements reorganize the catalytic centre of the protein to carry out two canonical reactions in a very non-canonical type of bifunctionality.**

Aldolases constitute a distinct group of lyase enzymes that catalyse the stereospecific addition of a nucleophilic donor substrate to an electrophilic acceptor<sup>7–9</sup>. Class I aldolases are commonly homotetramers<sup>10</sup> and are found in some bacteria, archaea and higher eukaryotes<sup>11,12</sup>. They activate their donor substrate by forming a Schiff base with the ε-amino group of a conserved lysine residue<sup>13</sup>. Aldolases of class II are found in bacteria and fungi and use divalent metal cations (mostly Zn<sup>2+</sup>, but frequently Fe<sup>2+</sup> or Co<sup>2+</sup>) to activate the donor nucleophile<sup>14</sup>. Although the two classes of aldolases do not show homologies in primary structure and are thus phylogenetically distinct, they are structurally related and belong to the family of (β/α)<sub>8</sub> TIM barrel enzymes<sup>15–18</sup>. Further members of this family include most other aldolases, such as the homodecameric class IA FBP aldolase<sup>4,19</sup>, transaldolase, deoxyribose phosphate aldolase, 2-keto-3-deoxy-(6-phospho)-gluconate (KD(P)G) aldolase and 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase<sup>20</sup>.

Fructose-1,6-bisphosphate aldolase (E.C. 4.1.2.13) catalyses the reversible aldol cleavage of FBP into DHAP and GAP in the Embden–Meyerhof–Parnas pathway<sup>21</sup> either in glycolysis or gluconeogenesis and in the Calvin–Benson cycle<sup>1</sup>. Most recently an enzyme previously described as an archaeal FBP phosphatase in *Thermococcus kodakarensis* and *Sulfolobus tokodaii*<sup>22</sup> was shown to be indeed a bifunctional FBP aldolase/phosphatase (FBPAP)<sup>2</sup>. Although this enzyme is physiologically unrelated to any known aldolase, it catalyses the reaction following the class I mechanism<sup>23</sup> involving a lysine Schiff base<sup>24,25</sup>. The aldolase reaction is a classic case for a fully reversible enzymatic

reaction<sup>9</sup>, but in FBPAP – in spite of a similar mechanism – the *K<sub>m</sub>* values for the aldol condensation and the aldol cleavage differ by a factor of 1,000, and the intrinsic phosphatase activity renders the process irreversible<sup>2</sup>. Orthologues of FBPAP are found in most genomes of archaea and early bacterial lineages.

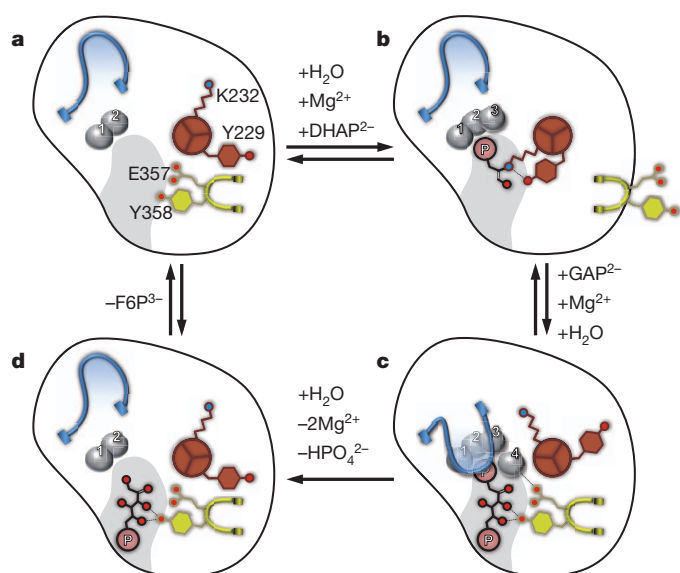
A structure of the enzyme from *S. tokodaii* revealed a novel tertiary structure with a ferredoxin-like fold in the amino (N)-terminal part and similarities to bacterial S-adenosyl methionine decarboxylases in the carboxy (C)-terminal part. The protein formed globular homooctamers and had four Mg<sup>2+</sup> ions and FBP bound to the active site<sup>20,22</sup>. To understand the twofold reactivity of FBPAP we have crystallized and characterized the enzyme from the hyperthermophilic crenarchaeon *T. neutrophilus* (*Tn*FBPAP) in a ligand-free state, with the substrates of the two reaction steps, DHAP and FBP, and with the product F6P. *Tn*FBPAP is highly similar to the *S. tokodaii* orthologue, differing predominantly in a C-terminal extension that embraces another protomer of the conserved octamer (Supplementary Fig. 1)<sup>22</sup>. However, three loop regions surrounding the active site of the enzyme substantially change their conformation when substrate, reaction intermediate or product is bound. The re-orientation of these loops during the catalytic cycle alters the structure and functionality of the active site fundamentally, while keeping the substrate locked in place. Unlike bi- or multifunctional enzymes described previously, FBPAP does not recruit distinct domains for its reactions or connect different active sites by substrate channelling. Instead the enzyme remodels its single active centre in order to bring the amino-acid side chains and cofactors—divalent Mg cations—into place.

In the following, the loop region from residues 220 to 235 will be termed the ‘aldolase loop’, as residues K232 and Y229 are essential for aldolase activity. It is observed in three different conformations that we designate ‘in’, ‘out’ and ‘locked’ (Fig. 1). The largest conformational changes were observed in the loop encompassing residues 89–111, the ‘phosphatase lid’ that serves to fix the intermediate FBP in the binding pocket for the phosphatase reaction. It is seen in two conformations, ‘open’ and ‘closed’. The third flexible loop is the ‘anchor loop’ from residues 353 to 364 that fixes FBP after the aldol condensation and attains an ‘in’ or an ‘out’ conformation. The consecutive binding of additional Mg<sup>2+</sup> ions to the protein is then key to switching its catalytic functionality.

In the substrate-free structure of *Tn*FBPAP, two Mg<sup>2+</sup> ions are bound to the protein, with Mg1 coordinated by residues D11, H18 and D52, and Mg2 by residues D52, D53, D132 and D234 from the aldolase loop. Both Mg<sup>2+</sup> ions show an octahedral coordination environment, completed by three water ligands for Mg1 and by two water ligands for Mg2 (Fig. 2a and Supplementary Figs 2 and 3a). In the structure, a substrate binding cleft is visible close to the ions, and four of the five water ligands are at the surface of the protein. The phosphatase lid is in its ‘open’ conformation, the aldolase loop is ‘out’ and the anchor loop is ‘in’ (Figs 1a and 2a and Supplementary Fig. 3a). A structural analysis of EDTA-treated *Tn*FBPAP shows that Mg1 can

<sup>1</sup>Lehrstuhl für Biochemie, Institut für organische Chemie und Biochemie, Albert-Ludwigs-Universität Freiburg, Albertstrasse 21, 79104 Freiburg, Germany. <sup>2</sup>Lehrstuhl für Mikrobiologie, Albert-Ludwigs-Universität Freiburg, Schänzlestrasse 1, 79104 Freiburg, Germany. <sup>3</sup>BIOSS Centre for Biological Signalling Studies, Albert-Ludwigs-Universität Freiburg, Hebelstrasse 25, 79104 Freiburg, Germany.

\*These authors contributed equally to this work.



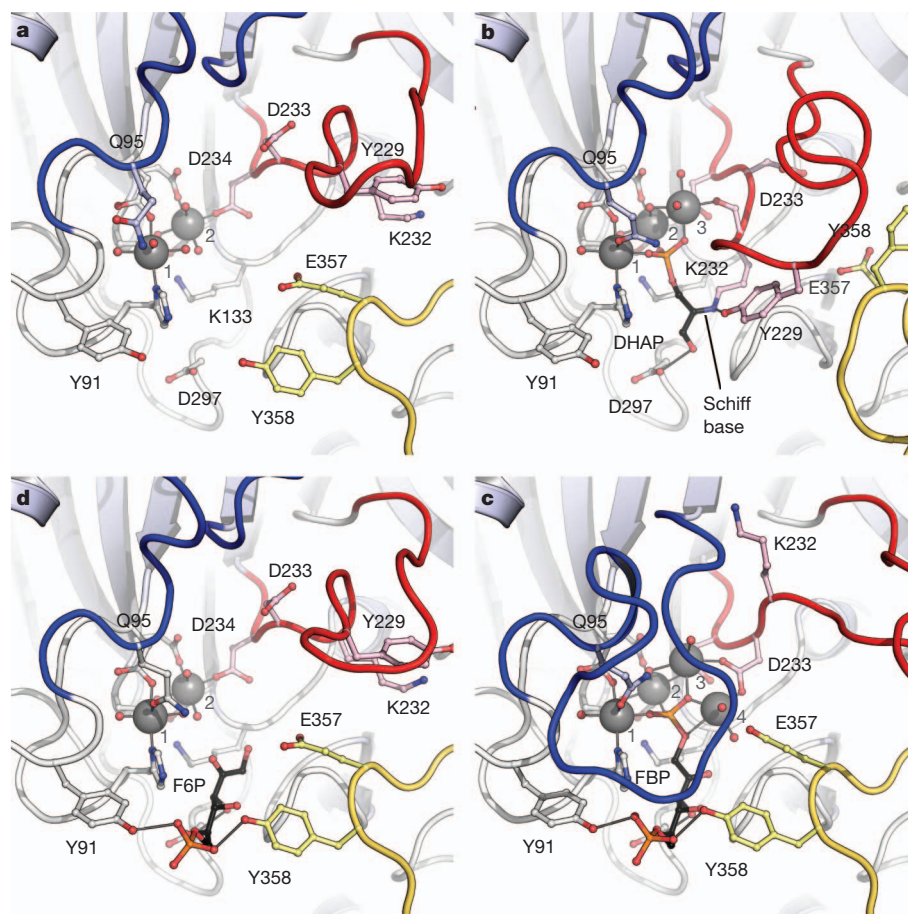
be removed, while Mg2 remains bound to the protein. In a calorimetric titration Mg1 binds with a  $K_d$  of 27  $\mu$ M (Supplementary Fig. 5), but no binding of further magnesium occurs. In this state the enzyme can bind its first substrate, DHAP (but not GAP), with a  $K_d$  of 25  $\mu$ M (Supplementary Fig. 7), thereby creating a further  $Mg^{2+}$  binding site.

Consequently, the crystal structure with bound DHAP contains three  $Mg^{2+}$ , and each of the C1-phosphoryl oxygen atoms coordinates

**Figure 1 | Representation of the reaction steps of FBP aldolase/phosphatase.** The phosphatase loop is shown in blue, the aldolase loop in red and the anchor loop in yellow. **a**, In the unliganded state the enzyme binds two  $Mg^{2+}$  ions and the active site is accessible. **b**, Binding of the substrate dihydroxyacetone phosphate is followed by a third  $Mg^{2+}$  ion that redirects the aldolase loop to form a Schiff base with the substrate. The anchor loop is retracted to accommodate this conformational change. **c**, The co-substrate glyceraldehyde-3-phosphate binds and the enzyme catalyses the aldol condensation. Release of the Schiff base leads to a rearrangement of the aldolase loop that creates a binding site for a fourth  $Mg^{2+}$  ion. The phosphatase loop changes its conformation drastically closing the active site. **d**, Upon hydrolysis of the phosphate group at C1, the tetranuclear Mg site disassembles and the three loops attain their original conformations.

one of the cations (Figs 1b and 2b and Supplementary Fig. 3b). At both Mg1 and Mg2, two phosphoryl oxygens replace the coordinated water molecules (Supplementary Figs 2 and 3b). In the next step in the reaction cascade, the backbone carbonyl oxygen of residue K232 occupies a remaining, free coordination site at the metal ion. The following major structural rearrangement switches the aldolase loop to the 'in' position and the anchor loop to the 'out' position. The phosphatase lid remains open, but one of its residues, Q95, rearranges to become a ligand to Mg1, likely stabilizing the complex metal site. More importantly, in the 'in' conformation of the aldolase loop K232 is in close proximity to the substrate and is able to form a protonated Schiff base intermediate with DHAP that is visible in the crystal structure (Figs 1b and 2b and Supplementary Fig. 3b).

The class-I aldolase mechanism requires a base that abstracts a proton in the next step<sup>26,27</sup>, and in *TnFBPAP* residue Y229 is positioned ideally to fulfil this role. It is part of the aldolase loop and swings

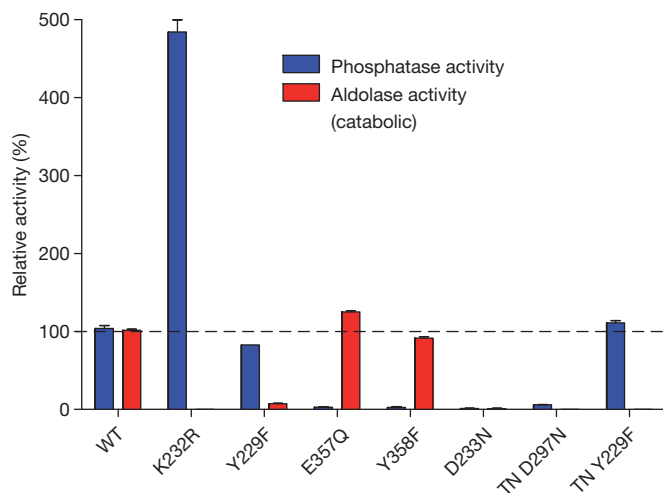


**Figure 2 | The successive reaction steps of FBP aldolase/phosphatase in the crystal structures.** **a**, The unliganded enzyme; **b**, the complex with dihydroxyacetone phosphate (DHAP); **c**, the complex with fructose-1,6-bisphosphate (FBP); **d**, the product complex with fructose-6-phosphate (F6P).

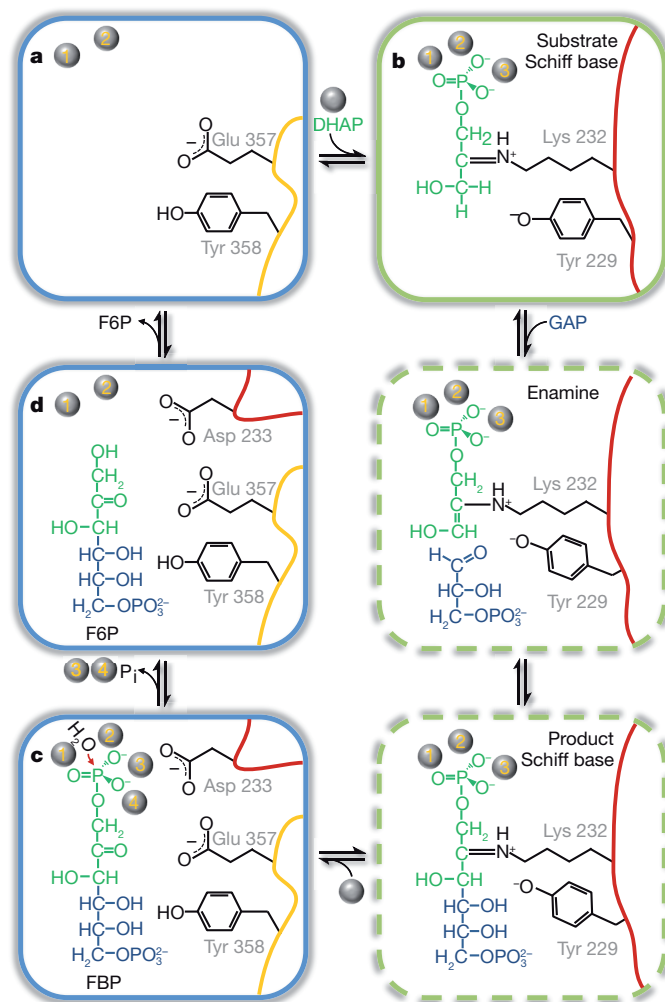
The reaction is guided by the successive binding of a third and fourth  $Mg^{2+}$  ion that orchestrate conformational rearrangements to position the aldolase (red), phosphatase (blue) and anchor loops (yellow) for their individual tasks.

into the active site when K232 forms the Schiff base. The distinct and isolated role of K232 and Y229 is emphasized by the fact that variants in these positions show only slightly reduced (Y229F) or even enhanced (K232R) phosphatase activity, whereas aldolase activity is completely abolished (Fig. 3 and Supplementary Table 2). The structure of the Y229F variant of FBPAP shows binding of DHAP with two  $Mg^{2+}$  ions, but the formation of the Schiff base is not observed. After its deprotonation, Y229 abstracts the *pro-S* proton at C3 of DHAP. This is ascertained by residue D297 that forms a short (2.6 Å) hydrogen bond with the hydroxyl group at C3 so that the *pro-S* H atom is fixed to face Y229 (Supplementary Fig. 3b). Proton abstraction results in a tautomeric rearrangement to yield an enamine intermediate. For the progression of the aldol condensation, the second substrate, GAP, then needs to bind to the active site. In the structure, the binding site for GAP is occupied by a second molecule of DHAP, giving a clear indication for the binding mode of the substrate, while being unreactive for the aldol condensation (Supplementary Fig. 3b).

C–C bond formation between the DHAP enamine and GAP yields the product Schiff base (Fig. 4) that is released to trigger the next conformational rearrangement in preparation for the phosphatase step. As the aldolase loop flips outward, the anchor loop changes back to the ‘in’ conformation, and one of its key residues, Y358, forms hydrogen bonds to the C4 OH group and the C6 phosphate of FBP (Figs 1c and 2c and Supplementary Fig. 3c). The phosphatase lid switches to ‘closed’, preventing access to the active site cavity. Concomitantly, the aldolase loop attains its ‘locked’ conformation, wherein K232 forms a hydrogen bond to the backbone carbonyl group of P111 at the base of the phosphatase loop (Supplementary Fig. 4). In effect, the position of the neighbouring D233 inverts, and its  $\beta$ -carboxy group becomes key to the creation of an additional binding site for  $Mg_4$ . Two water ligands at  $Mg_4$  are hydrogen-bonded to the side chain of E357 in the anchor loop and the four  $Mg^{2+}$  cations of this functional state are tightly grouped around the C1 phosphate of FBP (Supplementary Fig. 3b). The crystal structure, at a resolution of 1.3 Å, indeed shows a clear geometric distortion around the phosphorus atom, resulting in an O–P–O angle of approximately  $90^\circ$  for the two oxygen atoms coordinating  $Mg_4$  (Fig. 2c). Note that before the binding of  $Mg_4$  the three other  $Mg^{2+}$  ions are bound on the same side of the



**Figure 3 | Catalytic activity of distinct variants of FBP aldolase/phosphatase.** The two activities of the enzyme are tightly linked to distinct amino-acid residues and can be inactivated separately. Mutagenesis of K232 or Y229 on the aldolase loop abolishes aldolase activity while phosphatase activity is only slightly reduced (Y229F) or is even enhanced (K232R). In contrast, E357 and Y358 on the anchor loop are essential for phosphatase activity, but alterations here (E357Q, Y358F) do not affect aldolase activity. A D233N mutant protein was impaired in both aldolase and phosphatase activity. WT, wild type.



**Figure 4 | Proposed reaction mechanism of FBP aldolase/phosphatase.** The steps of the aldol condensation reaction are boxed in green, those of the phosphatase reaction in blue. Crystal structures are available for the intermediates in solid boxed, whereby the lettering follows Figs 1 and 2. The aldolase loop is indicated by a red line, the anchor loop by a yellow line.

phosphate group (Fig. 2b), thereby counteracting the formation of the trigonal bipyramidal transition state of phosphate ester hydrolysis. Variants of the anchor loop residues E357Q and Y358F exhibit unaltered aldolase activity, whereas phosphatase activity is fully abolished (Fig. 3). As described for the *S. tokodaii* enzyme, the structure allows the identification of a water molecule coordinated by  $Mg_2$  and  $Mg_3$  that conducts an in-line  $S_N2$  nucleophilic attack onto the phosphorus atom (Supplementary Fig. 2). This leads to an inversion of the ligand environment and the cleavage of the phosphoester bond to the sugar substrate.

At this point, the particular mechanistic intricacy of *Tn*FBPAP becomes apparent. After binding of DHAP, the aldolase and phosphatase reaction steps were supported by the subsequent addition of tightly coordinated  $Mg^{2+}$  ions. The metal triad or tetrad was organized and stabilized by the phosphate group at its centre, but as this phosphate inverts its geometry upon hydrolysis it no longer fits the arrangement of the surrounding cations. The significant free enthalpy of the hydrolysis reaction provides the driving force to overcome the stability of the complex and disassemble the centre. These changes lead back to the structure of the complex with the product F6P in the open-chain form. This structure contains only two  $Mg^{2+}$  ions and shows all three loops in the conformations observed in the ligand-free form of the enzyme (Figs 1d and 2d and Supplementary Fig. 3d). Inorganic phosphate itself was not found to bind in crystal soaking experiments.

*TnFBPAP* catalyses a multi-step reaction by remodelling its active site according to the respective catalytic requirements (Fig. 4). At first glance, this seems more elaborate than to combine different enzymatic modules into a multi-enzyme complex as seen, for example, in pyruvate dehydrogenase<sup>28,29</sup>. Yet *TnFBPAP* likely represents an ancestral gluconeogenic enzyme and the functional optimizations observed are an adaptation to high-temperature environments, where the instability of the substrates DHAP and GAP presents a serious problem<sup>2</sup>. Large conformational changes and the coupling to the exergonic phosphate hydrolysis allow the enzyme to render the aldolase reaction irreversible to assure that the cellular levels of the sensitive triose phosphates can be kept low. Besides its evolutionary impact, this enzyme sets an elucidatory example for how consecutive dramatic conformational changes can reorganize an active centre to perform two drastically different catalytic steps in a highly controlled and ordered sequence.

## METHODS SUMMARY

C-terminally His<sub>6</sub>-tagged *TnFBPAP* (Tneu\_0133) was heterologously produced in *Escherichia coli* as described previously<sup>2</sup>. Crystals were grown at 20 °C by the hanging-drop vapour-diffusion method. Two microlitres of protein solution were mixed with 2 µl of a reservoir solution containing 8% (w/v) of polyethylene glycol 3350 and 0.1 M HEPES/NaOH at pH 7.0–8.0. The drops were equilibrated against reservoir solution without added protein. The crystals were briefly soaked in 15% (v/v) of 2R,3R-butane diol for cryoprotection before being plunged into liquid nitrogen. For the preparation of complexes with the compounds dihydroxyacetone phosphate, fructose-1,6-bisphosphate and fructose-6-phosphate, the native crystals were soaked in 15% (v/v) of 2R,3R-butane diol and 100 mM of the respective substrate for 20 min before being treated for cryo-protection as described above. Diffraction data were collected at beamline X06SA at the Swiss Light Source (Villigen, Switzerland) at an X-ray wavelength of 1.0 Å. The native structure was refined to  $R_{\text{cryst}} = 0.165$  and  $R_{\text{free}} = 0.188$  at a resolution of 1.52 Å. The crystal structures in complex with DHAP, FBP and F6P were solved by molecular replacement with the native structure as the search model (Supplementary Table 1). Activity assays for the *Cenarchaeum symbiosum* and *T. neutrophilus* enzymes were performed at 40 °C and 48 °C, respectively, using a coupled spectrophotometric assay as described previously<sup>2</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 2 March; accepted 13 August 2011.**

**Published online 9 October 2011.**

1. Berg, I. A. *et al.* Autotrophic carbon fixation in archaea. *Nature Rev. Microbiol.* **8**, 447–460 (2010).
2. Say, R. F. & Fuchs, G. Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* **464**, 1077–1081 (2010).
3. Stetter, K. O. Hyperthermophiles in the history of life. *Phil. Trans. R. Soc. B* **361**, 1837–1842 (2006).
4. Siebers, B. *et al.* Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *J. Biol. Chem.* **276**, 28710–28718 (2001).
5. Messner, P., Pum, D., Sara, M., Stetter, K. O. & Sleytr, U. B. Ultrastructure of the cell envelope of the archaeobacteria *Thermoproteus tenax* and *Thermoproteus neutrophilus*. *J. Bacteriol.* **166**, 1046–1054 (1986).
6. Zillig, W. *et al.* The phylogenetic relations of DNA-dependent RNA polymerases of archaeobacteria, eukaryotes, and eubacteria. *Can. J. Microbiol.* **35**, 73–80 (1989).
7. Alefounder, P. R., Baldwin, S. A., Perham, R. N. & Short, N. J. Cloning, sequence analysis and over-expression of the gene for the class II fructose 1,6-bisphosphate aldolase of *Escherichia coli*. *Biochem. J.* **257**, 529–534 (1989).
8. Fothergill-Gilmore, L. A. & Michels, P. A. Evolution of glycolysis. *Prog. Biophys. Mol. Biol.* **59**, 105–235 (1993).
9. Rutter, W. J. Evolution of aldolase. *Fed. Proc.* **23**, 1248–1257 (1964).

10. Lorentzen, E., Siebers, B., Hensel, R. & Pohl, E. Structure, function and evolution of the archaeal class I fructose-1,6-bisphosphate aldolase. *Biochem. Soc. Trans.* **32**, 259–263 (2004).
11. Penhoet, E., Kochman, M., Valentine, R. & Rutter, W. J. The subunit structure of mammalian fructose diphosphate aldolase. *Biochemistry* **6**, 2940–2949 (1967).
12. Tolan, D. R., Niclas, J., Bruce, B. D. & Lebo, R. V. Evolutionary implications of the human aldolase-A, -B, -C, and -pseudogene chromosome locations. *Am. J. Hum. Genet.* **41**, 907–924 (1987).
13. Lebherz, H. G. & Rutter, W. J. Distribution of fructose diphosphate aldolase variants in biological systems. *Biochemistry* **8**, 109–121 (1969).
14. Cooper, S. J. *et al.* The crystal structure of a class II fructose-1,6-bisphosphate aldolase shows a novel binuclear metal-binding active site embedded in a familiar fold. *Structure* **4**, 1303–1315 (1996).
15. Lorentzen, E. *et al.* Crystal structure of an archaeal class I aldolase and the evolution of (β $\alpha$ )<sub>8</sub> barrel proteins. *J. Biol. Chem.* **278**, 47253–47260 (2003).
16. Hester, G. *et al.* The crystal structure of fructose-1,6-bisphosphate aldolase from *Drosophila melanogaster* at 2.5 Å resolution. *FEBS Lett.* **292**, 237–242 (1991).
17. Gambin, S. J. *et al.* The crystal structure of human muscle aldolase at 3.0 Å resolution. *FEBS Lett.* **262**, 282–286 (1990).
18. Blom, N. S., Tetreault, S., Coulombe, R. & Sygusch, J. Novel active site in *Escherichia coli* fructose-1,6-bisphosphate aldolase. *Nature Struct. Biol.* **3**, 856–862 (1996).
19. Imanaka, H., Fukui, T., Atomi, H. & Imanaka, T. Gene cloning and characterization of fructose-1,6-bisphosphate aldolase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Biosci. Bioeng.* **94**, 237–243 (2002).
20. Andreeva, A. *et al.* Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.* **36**, 419–425 (2008).
21. Ronimus, R. S. & Morgan, H. W. Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea* **1**, 199–221 (2003).
22. Nishimasu, H., Fushinobu, S., Shoun, H. & Wakagi, T. The first crystal structure of the novel class of fructose-1,6-bisphosphatase present in thermophilic archaea. *Structure* **12**, 949–959 (2004).
23. St-Jean, M., Blonski, C. & Sygusch, J. Charge stabilization and entropy reduction of central lysine residues in fructose-bisphosphate aldolase. *Biochemistry* **48**, 4528–4537 (2009).
24. Grazi, E., Rowley, P. T., Cheng, T., Tchola, O. & Horecker, B. L. The mechanism of action of aldolases. III. Schiff base formation with lysine. *Biochem. Biophys. Res. Commun.* **9**, 38–43 (1962).
25. Rose, I. A. & Rieder, S. V. Studies on the mechanism on the aldolase reaction; isotope exchange reactions of muscle and yeast aldolase. *J. Biol. Chem.* **231**, 315–329 (1958).
26. Lai, C. Y., Tchola, O., Cheng, T. & Horecker, B. L. The mechanism of action of aldolases. 8. The number of combining sites in fructose diphosphate aldolase. *J. Biol. Chem.* **240**, 1347–1350 (1965).
27. St-Jean, M., Lafrance-Vanasse, J., Liotard, B. & Sygusch, J. High resolution reaction intermediates of rabbit muscle fructose-1,6-bisphosphate aldolase: substrate cleavage and induced fit. *J. Biol. Chem.* **280**, 27262–27270 (2005).
28. Ciszak, E. M., Korotchkina, L. G., Dominiak, P. M., Sidhu, S. & Patel, M. S. Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase. *J. Biol. Chem.* **278**, 21240–21246 (2003).
29. Perham, R. N. Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* **69**, 961–1004 (2000).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by Deutsche Forschungsgemeinschaft (grant Ei520/3-2 to O.E., Fu118/15-4 and 15-5 to G.F.). Diffraction data were collected at beam lines X06SA and X06DA at the Swiss Light Source (Villigen, Switzerland). The authors thank the beam line staff for assistance during data collection.

**Author Contributions** J.D. and W.L. crystallized the protein and collected diffraction data, J.D., W.L. and O.E. built and refined the structural models, R.F.S. created and analysed the variant proteins, J.D., R.F.S., W.L., G.F. and O.E. designed the experiments, O.E. wrote the manuscript.

**Author Information** Atomic coordinates and structure factors for the reported crystal structures are deposited in Protein Data Bank under accession numbers 3T2B (ligand-free), 3T2C (DHAP-bound), 3T2D (FBP-bound), 3T2E (F6P-bound), 3T2F (EDTA-soak with DHAP) and 3T2G (Y229F variant with DHAP). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to O.E. ([einsle@biochemie.uni-freiburg.de](mailto:einsle@biochemie.uni-freiburg.de)).

## METHODS

**Cloning, expression and purification.** C-terminally His<sub>6</sub>-tagged *TnFBPAP* (Tneu\_0133) was heterologously produced in *E. coli* as described previously<sup>2</sup>. Frozen *E. coli* cells (30 g wet mass) were suspended in 30 ml re-suspension buffer (20 mM Tris/HCl pH 7.8, 20 mM MgCl<sub>2</sub>, 0.5 mM phenylmethanesulphonyl fluoride) containing 0.1 mg ml<sup>-1</sup> DNase I and passed twice through a French pressure cell at 137 MPa. The cell lysate was centrifuged at 150,000g for 1 h (4 °C). The cell extract (10 ml) was loaded onto a 1 ml Ni<sup>2+</sup>-chelating Sepharose affinity column (GE Healthcare) equilibrated with running buffer (20 mM Tris/HCl pH 7.8, 50 mM MgCl<sub>2</sub> and 200 mM NaCl) at a flow rate of 1 ml min<sup>-1</sup>. The column was washed with running buffer containing 70 mM imidazole and developed with a linear gradient of imidazole. Recombinant *TnFBPAP* eluted at a concentration of 275 mM imidazole. Active fractions were pooled and the buffer was exchanged for 50 mM 3-(N-morpholino)propanesulphonate (MOPS) buffer at pH 8.0 with 10 mM MgCl<sub>2</sub> using a PD-10 column (GE Healthcare). For subsequent heat precipitation, the enzyme was incubated for 30 min at 80 °C, cooled on ice for 15 min and centrifuged at 17,000g for 30 min (4 °C). For crystallization, the supernatant with purified protein was concentrated by ultrafiltration to a final concentration of 7 mg ml<sup>-1</sup>.

**Crystallization.** Crystals were grown at 20 °C by the hanging-drop vapour-diffusion method. Two microlitres of protein solution were mixed with 2 µl of a reservoir solution containing 8% (w/v) of polyethylene glycol 3350 and 0.1 M HEPES/NaOH at pH 7.0–8.0. The drops were equilibrated against reservoir solution without added protein. Native crystals appeared after 1 h and reached their maximum size after approximately 3 days. The crystals were briefly soaked in 15% (v/v) of 2R,3R-butane diol for cryoprotection before being plunged into liquid nitrogen. For the preparation of complexes with the compounds dihydroxyacetone phosphate, fructose-1,6-bisphosphate and fructose-6-phosphate, the native crystals were soaked in 15% (v/v) of 2R,3R-butane diol and 100 mM of the respective substrate for 20 min before being treated for cryo-protection as described above.

**Data collection and structure determination.** Diffraction data were collected at beamline X06SA at the Swiss Light Source (Villigen, Switzerland) at an X-ray wavelength of 1.0 Å. The crystals of *TnFBPase* belonged to the tetragonal space group *I*422 with one monomer per asymmetric unit. Data were indexed and integrated using MOSFLM<sup>30</sup> and scaled with sCALA<sup>31</sup>. Structure solution was carried out by molecular replacement using the program MOLREP<sup>32</sup> and the structure of the homologous enzyme from *S. tokodaii* (PDB-ID 1UMG) as the initial search model. Refinement was performed with REFMAC5 (ref. 33) and model building was performed using coot<sup>34</sup>. The native structure was refined to  $R_{\text{cryst}} = 0.165$  and  $R_{\text{free}} = 0.188$  (ref. 35) at a resolution of 1.52 Å. The crystal

structures in complex with DHAP, FBP and F6P were solved by molecular replacement with the native structure as the search model (Supplementary Table 1).

**Site-directed mutagenesis.** Mutagenesis was performed on the synthetic *C. symbiosum* gene<sup>1</sup>. The first mutations were introduced into the expression vector pT7-7 (ref. 36) carrying the gene (FBP\_C.symb-X-pT7-7 (ref. 2)) by reverse PCR using a single mutagenic oligonucleotide<sup>9</sup>. PCR conditions were as follows: 25 cycles of 20 s denaturation at 98 °C, 20 s primer annealing, and elongation at 72 °C using Phusion DNA polymerase (New England Biolabs). The PCR products were incubated at 37 °C with 20 U of *DpnI* for 3 h to digest the methylated template plasmid. After amplification of the plasmid in *E. coli* DH5 $\alpha$ , the mutation was confirmed by sequencing. Competent *E. coli* BL21 (DE3) Rosetta2 cells (Novagen) were transformed with the corresponding plasmid and grown at 37 °C in 1-l flasks with self-inducing medium<sup>37</sup> containing 100 µg ml<sup>-1</sup> ampicillin and 34 µg ml<sup>-1</sup> chloramphenicol. After 5–6 h at 37 °C, the temperature was lowered to 20 °C and the culture was grown for 14 h at 20 °C. Cells were collected by centrifugation and stored in liquid nitrogen until further processing.

**Production of FBP aldolase/phosphatase mutants from *C. symbiosum*.** Preparation of cell extract from 3–6 g wet mass of frozen *E. coli* cells and affinity chromatography were performed following the procedure described above for the *T. neutrophilus* protein.

**FBP aldolase/phosphatase enzyme assays.** Activity assays for the *C. symbiosum* and *T. neutrophilus* enzymes were performed at 40 °C and 48 °C, respectively, using a coupled spectrophotometric assay as described previously<sup>2</sup>.

30. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* No. 26, (1992).
31. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
32. Vagin, A. A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
33. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
34. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
35. Brünger, A. T. Assessment of phase accuracy by cross validation - the free R-value - methods and applications. *Acta Crystallogr. D* **49**, 24–36 (1993).
36. Tabor, S. & Richardson, C. C. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl Acad. Sci. USA* **82**, 1074–1078 (1985).
37. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).

# Structural basis for the bifunctionality of fructose-1,6-bisphosphate aldolase/phosphatase

Shinya Fushinobu<sup>1\*</sup>, Hiroshi Nishimasu<sup>2\*</sup>, Daiki Hattori<sup>1</sup>, Hyun-Jin Song<sup>1</sup> & Takayoshi Wakagi<sup>1</sup>

Enzymes catalyse specific reactions and are essential for maintaining life. Although some are referred to as being bifunctional, they consist of either two distinct catalytic domains or a single domain that displays promiscuous substrate specificity<sup>1</sup>. Thus, one enzyme active site is generally responsible for one biochemical reaction. In contrast to this conventional concept, archaeal fructose-1,6-bisphosphate (FBP) aldolase/phosphatase (FBPA/P) consists of a single catalytic domain, but catalyses two chemically distinct reactions of gluconeogenesis: (1) the reversible aldol condensation of dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (GA3P) to FBP; (2) the dephosphorylation of FBP to fructose-6-phosphate (F6P)<sup>2</sup>. Thus, FBPA/P is fundamentally different from ordinary enzymes whose active sites are responsible for a specific reaction. However, the molecular mechanism by which FBPA/P achieves its unusual bifunctionality remains unknown. Here we report the crystal structure of FBPA/P at 1.5-Å resolution in the aldolase form, where a critical lysine residue forms a Schiff base with DHAP. A structural comparison of the aldolase form with a previously determined phosphatase form<sup>3</sup> revealed a dramatic conformational change in the active site, demonstrating that FBPA/P metamorphoses its active-site architecture to exhibit dual activities. Thus, our findings expand the conventional concept that one enzyme catalyses one biochemical reaction.

FBPA/P was initially identified as class V fructose-1,6-bisphosphatase (FBPase)<sup>4</sup>, and is responsible for gluconeogenesis in the hyperthermophilic archaeon *Thermococcus kodakaraensis*<sup>5</sup>. Class V FBPases lack sequence homology with other FBPase or FBP aldolase (FBPA) proteins, and are found in the genomes of virtually all Archaea and deeply branching Bacteria (Supplementary Fig. 1). FBPA/P shows higher activity in the condensation reaction than in the FBP cleavage reaction<sup>2</sup> (Supplementary Fig. 2), indicating its involvement in gluconeogenesis, rather than in glycolysis. Thus, FBPA/P has been proposed to represent an ancestral gluconeogenic enzyme that ensures unidirectional gluconeogenesis in chemolithoautotrophic organisms<sup>2</sup>. We previously determined the crystal structure of an FBPA/P protein from *Sulfolobus tokodaii* (ST0318) at 1.8-Å resolution, in complex with FBP and catalytically essential Mg<sup>2+</sup> ions, which represents the 'FBPase form'<sup>3</sup>. The structure revealed a unique fold and lacked similarity to those of other known FBPase and FBPA proteins. Unexpectedly, biochemical experiments indicated that a conserved lysine residue (Lys 232 of ST0318), which is located away from the bound FBP molecule in the ST0318 structure, forms a Schiff base with DHAP during the FBPA reaction<sup>2</sup>, suggesting that FBPA/P undergoes a large structural change for its dual activities. However, the molecular mechanism by which FBPA/P exhibits its bifunctionality remains unknown. Here, we present the crystal structure of ST0318 in the DHAP-Schiff base intermediate state, which represents the 'FBPA form'. A comparison of the high-resolution structures of FBPA/P in the aldolase and phosphatase forms provides the structural basis for its bifunctionality.

We measured the FBPA and FBPase activities of the purified recombinant ST0318 enzyme, and confirmed that the enzyme exhibits both activities (Table 1), demonstrating that ST0318 is also an FBPA/P enzyme. We crystallized ST0318 in the presence of DHAP and Mg<sup>2+</sup>, and determined the crystal structure at 1.5-Å resolution (Fig. 1a, Supplementary Fig. 3a and Supplementary Table 1). The crystal belongs to the space group I422, and is isomorphous to the previously determined FBPase form of ST0318 (PDB code 1UMG). The monomer structures are almost identical in the FBPA and FBPase forms, with root mean square deviations of 3.28 and 0.29 Å for the Cα atoms of the overall polypeptide and the regions without the three mobile loops (residues 97–110, 219–233 and 346–361). As in the FBPase form<sup>3</sup>, the protomers in the asymmetric units are related by crystallographic symmetry to form an octamer (Supplementary Fig. 4). Thus, the ternary and quaternary structures are virtually the same in these two forms. We observed an electron density for the bound DHAP in the active site, located at the dimer interface of the ring-shaped tetramer (Supplementary Fig. 5). As previously postulated<sup>2</sup>, Lys 232 formed a Schiff base intermediate with DHAP (Fig. 1a and Supplementary Fig. 3a). The Lys 232 Nζ atom and the DHAP C1, C2 and C3 atoms adopt a nearly planar arrangement, indicating that the intermediate is similar to the imine (iminium) form observed in a typical class I FBPA, for example rabbit muscle FBPA (rmFBPA)<sup>6</sup>. The intermediate may have been trapped, owing to the absence of the second substrate, GA3P. The DHAP phosphate group coordinates three Mg<sup>2+</sup> ions (Mg2–Mg4) (Supplementary Fig. 6), whereas the FBP 1-phosphate group coordinates four Mg<sup>2+</sup> ions (Mg1–Mg4) in the FBPase form<sup>3</sup>. The DHAP hydroxyl group is recognized by Arg 266 and Asp 287 through direct hydrogen bonds, and by Gln 242' (the prime symbol indicates residues from the neighbouring protomer) through a water-mediated hydrogen bond. These hydrogen-bonding interactions are likely to be critical for the discrimination between DHAP and GA3P, as GA3P has a carbonyl group and lacks a hydroxyl

**Table 1 | Activities of wild type and mutants of ST0318**

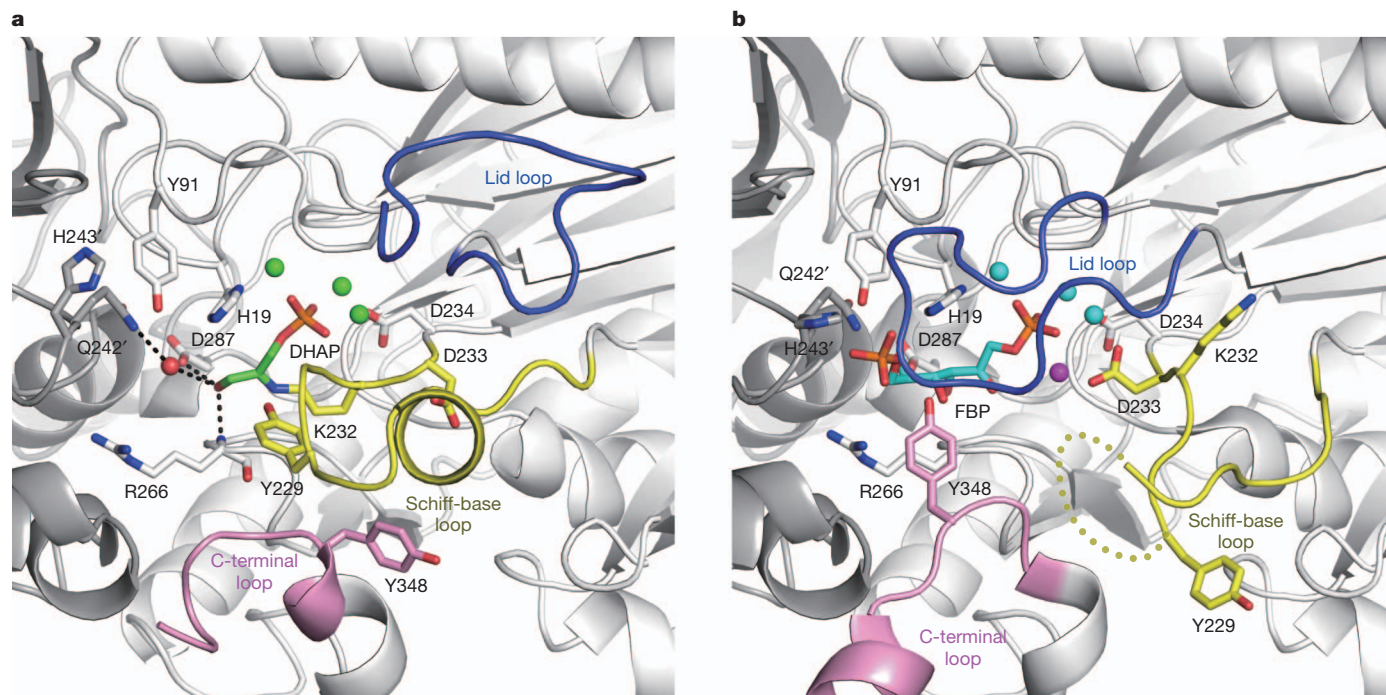
Enzyme	$k_{\text{cat}}$ (s <sup>-1</sup> )	$K_{\text{m}}$ (mM)	$k_{\text{cat}}/K_{\text{m}}$ (s <sup>-1</sup> mM <sup>-1</sup> )
FBPase			
Wild type	0.62 ± 0.02	0.027 ± 0.003	23
Y229F	0.66 ± 0.02	0.027 ± 0.003	25
Y348F	0.26 ± 0.01	0.036 ± 0.002	7.2
FBPA (anabolic direction)			
Wild type	0.91 ± 0.04	0.19 ± 0.02	4.7
Y229F	ND*	–	–
Y348F	0.10 ± 0.01	0.34 ± 0.04	0.29
FBPA (catabolic direction)			
Wild type	0.027 ± 0.011†	–	–
Y229F	ND*	–	–
Y348F	0.026 ± 0.004†	–	–

\* ND, not detected.

† Enzyme activity was measured with 5 mM FBP.

<sup>1</sup>Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. <sup>2</sup>Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan.

\*These authors contributed equally to this work.



**Figure 1 | Active sites of ST0318 in the two forms.** **a**, **b**, DHAP–Schiff base complex (the FBPA form) (**a**) and FBP complex (the FBPase form, PDB code 1UMG)<sup>3</sup> (**b**). DHAP and FBP are coloured green and cyan, respectively. The neighbouring protomer is coloured dark grey, and its residues are labelled with prime symbols. The lid, Schiff-base and C-terminal loops are coloured blue,

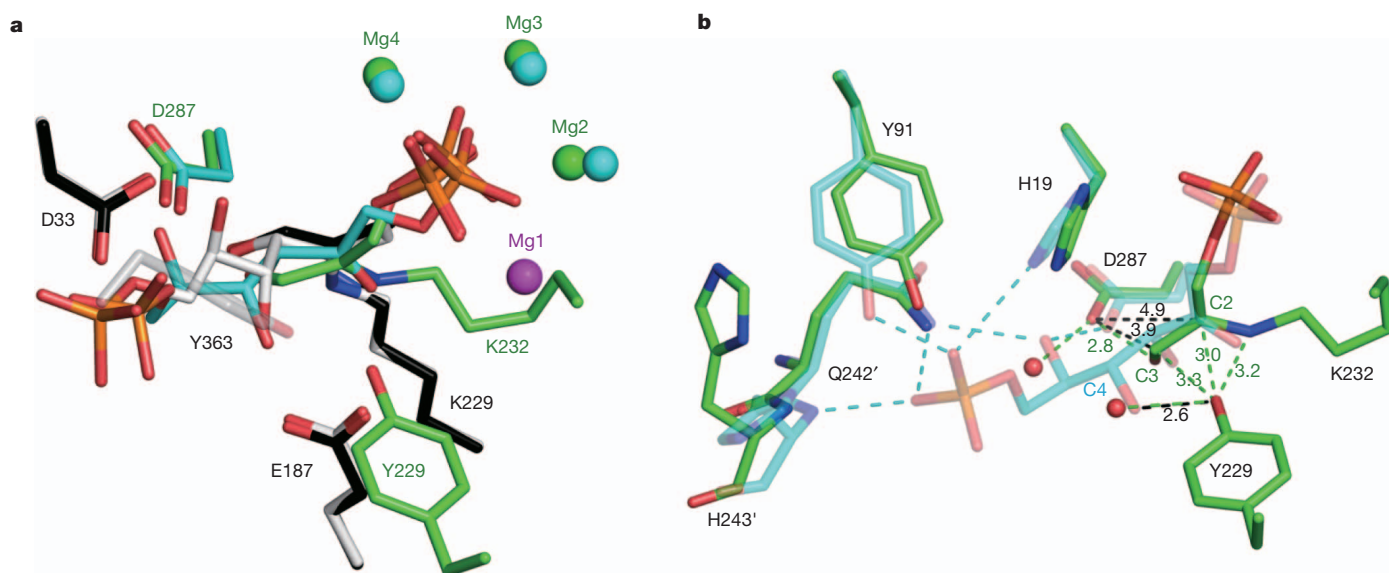
yellow and pink, respectively. In (**a**), Mg<sup>2+</sup> ions (Mg2–4) and a water molecule are shown as green and red spheres, respectively. Hydrogen bonds are shown as black dashed lines. In (**b**), Mg<sup>2+</sup> ions (Mg1 and Mg2–4) are shown as magenta and cyan spheres, respectively. The disordered region is indicated by a yellow dashed line.

group at the C3 position. The enzyme displayed the activity under the crystallization conditions (albeit approximately 50% compared with that under standard conditions), and the dissolved crystals exhibited activity comparable to that of the purified enzyme, confirming that the FBPA/P–DHAP complex structure determined here represents a genuine reaction intermediate. The DHAP molecule is located at a position similar to that of the FBP molecule in the FBPase form, demonstrating that FBPA/P catalyses both the FBPA and FBPase reactions at a single site.

A structural comparison of the FBPA and FBPase forms reveals striking differences in three loop regions at the active site: a lid loop (residues 97–110), a Schiff-base loop (residues 219–233) that contains the Schiff base-forming Lys232 residue, and a carboxy (C)-terminal loop (residues 346–361) (Fig. 1). Although we could confidently trace the polypeptide chain, except for part of the C-terminal loop (residues 354–360), the loop regions exhibit relatively higher average B-factors than that for the overall polypeptide (Supplementary Table 2), indicating their flexible nature. In the FBPase form, the lid and C-terminal loops interact with FBP, whereas the Schiff-base loop is further away from the active site and does not contact FBP<sup>3</sup> (Fig. 1b and Supplementary Fig. 3b). In the FBPA form, the lid and C-terminal loops are displaced outwards from the active site and do not contact DHAP, whereas the Schiff-base loop cuts into the active site, allowing Tyr 229 and Lys 232 to interact with DHAP (Fig. 1a and Supplementary Fig. 3a). Although Asp 234 similarly holds Mg2 in the two forms, Asp 233, which holds the catalytically essential Mg1 in the FBPase form<sup>3</sup>, is flipped out in the FBPA form, resulting in the dissociation of Mg1.

A structural comparison of ST0318 in the two forms with rmFBPA in the DHAP- and FBP–Schiff base intermediate states<sup>6,7</sup> reveals an unexpected similarity in the active-site configurations between FBPA/P and class I FBPA (Fig. 2a and Supplementary Fig. 7a). Notably, the covalent DHAP intermediates of ST0318 and rmFBPA superimpose well, even though the C $\alpha$  atoms of their Schiff base-forming lysine residues are far apart, owing to their distinct overall structures

(Supplementary Fig. 8). Based on this analogy, we propose a mechanism for FBPA/P-catalysed reactions (Fig. 3). The aldol condensation reaction consists of dehydration, carbanion formation, C3–C4 bond formation and hydration steps, in which general acid/base catalysts participate. In rmFBPA, Glu 187 reportedly functions as the general acid/base residue in multiple steps<sup>7</sup>. In addition, Asp 33 (ref. 8) or Tyr 363 (ref. 6) may be involved in the second carbanion formation step, by accepting a proton from the C3 methylene of the DHAP–Schiff base intermediate. Tyr 229 and Asp 287 in ST0318 are located at positions similar to those of Glu 187 and Asp 33 in rmFBPA (Fig. 2a and Supplementary Fig. 7a). Tyr 229 is located in the vicinity of the Lys 232 N $\zeta$  (3.2 Å), the DHAP C2 (3.0 Å) and C3 (3.3 Å), and the superposed FBP O4 (2.6 Å) atoms, whereas Asp 287 is further away from the DHAP C2 (4.9 Å) and C3 (3.9 Å) atoms (Fig. 2b and Supplementary Fig. 7b). Moreover, Tyr 229 is completely conserved in the FBPA/P proteins, whereas Asp 287 is replaced with a cysteine residue in the *Thermus thermophilus* and *Cenarchaeum symbiosum* FBPA/Ps (Supplementary Fig. 1). Indeed, the Y229F mutant retained the FBPase activity, but lost the FBPA activity (Table 1). These observations suggest that Tyr 229 serves as the catalytic acid/base residue for all steps (Fig. 3). In contrast to the FBPase form, the active site is accessible to the solvent in the FBPA form, and thus seems suitable for the hydration reaction. A comparison of the high-resolution structures of the two forms provides insight into the GA3P recognition mechanism. In the FBPA form, two water molecules form hydrogen bonds with Tyr 229 and Asp 287, and they are located at positions equivalent to the FBP 4- and 5-hydroxyl groups in the FBPase form, respectively (Fig. 2b and Supplementary Fig. 7b). In addition, the residues involved in the recognition of the FBP 6-phosphate group (His 19, Tyr 91, Gln 242' and His 243') are located at similar positions in the two forms. Thus, it is likely that the GA3P phosphate group is recognized by His 19, Tyr 91, Gln 242' and His 243', and the GA3P carbonyl and hydroxyl groups are recognized by Tyr 229 and Asp 287, respectively. The nucleophilic water molecule for the FBPase reaction is already bound to the FBPA form, and is held by Mg2 and Mg3

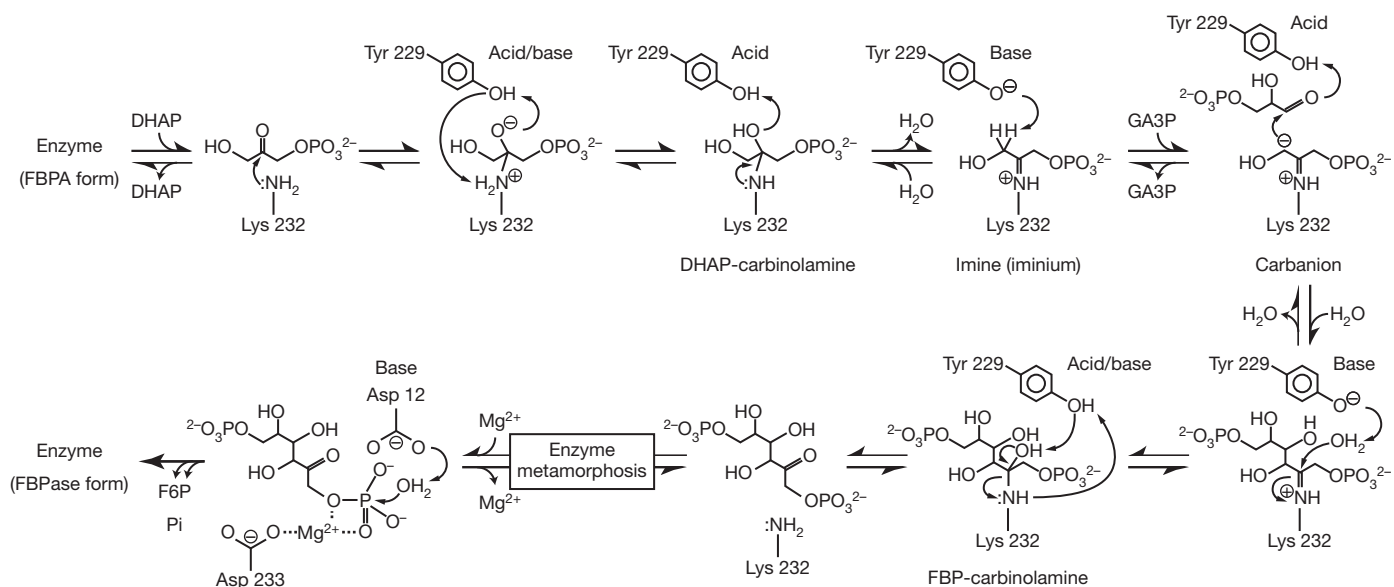


**Figure 2 | Catalytic components of FBPA.** **a**, Superimposition of the FBPA (green) and FBPAse forms (cyan, PDB code 1UMG) of ST0318, and the DHAP-Schiff base (enamine) intermediate (black, PDB code 2QUT) and the FBP-Schiff base intermediate (pericyclic transition state) of wild-type rmFBPA (white, PDB code 1ZAI), and Tyr 363 of the rmFBPA K146M mutant (semi-transparent grey, PDB code 2QUU). The residues of ST0318 and rmFBPA are labelled with green and black characters, respectively. The  $Mg^{2+}$  ion (Mg1) in the FBPAse form is shown as a magenta sphere. **b**, Superimposition of the FBPA

(green) and FBPAse forms (semi-transparent cyan, PDB code 1UMG) of ST0318. Water molecules in the FBPA form are shown as red spheres. The interactions in the FBPA and FBPAse forms are depicted by green and cyan dashed lines, respectively, and the distances (in Ångströms) between the protein atoms (Tyr 229 or Asp 287) and the DHAP-Schiff base are shown (green). The distances (in Ångströms) between Asp 287 and the DHAP C2 and C3 atoms and between Tyr 229 and the FBP O4 atom are also shown with black dashed lines.

(Supplementary Fig. 6). In addition, the catalytic base for the FBPAse reaction, Asp 12, is located at a similar position in both forms. However, DHAP is protected from dephosphorylation, probably because of the lack of Mg1, which is required for stabilizing the developing negative charge after phosphate cleavage in the FBPAse reaction<sup>3</sup>. Hydrolysis of the Schiff-base releases the loop containing Lys 232 from the active site, thereby enabling the enzyme to bind Mg1. In addition, the closure of the lid and C-terminal loops would be important for the FBPAse reaction, as they stabilize FBP binding. Gly 104 and Asn 105 on the lid loop interact with the FBP 6- and 1-phosphate groups, respectively, whereas Tyr 348 on the C-terminal loop interacts with the FBP 4-hydroxyl and 6-phosphate groups<sup>3</sup>.

In summary, the crystal structures of FBPA/P in the two forms revealed that FBPA/P achieves its bifunctionality by transforming its active-site architecture, through the toggle switch-like motions of the three mobile loop regions. Conformational fluctuations, for example loop movements and domain motions, are essential for substrate binding and product release in enzyme functions<sup>9–11</sup>. A recent study showed that a His/Trp biosynthesis isomerase, PriA, exhibits bisubstrate specificity through a substrate-induced metamorphosis of the active-site architecture<sup>12</sup>. However, the role of conformational fluctuations in enzyme catalysis and bifunctionality has been less clear. To our knowledge, this study is the first to elucidate the molecular mechanism by which an enzyme catalyses multiple chemical reactions at a single



**Figure 3 | Proposed mechanism of the FBPA and FBPAse reactions catalysed by ST0318.** The crystal structure determined here represents the imine (iminium) intermediate.

site. Furthermore, our findings suggest the possible existence of undiscovered enzymes that also catalyse multiple chemical reactions at a single site.

## METHODS SUMMARY

The wild-type and mutant ST0318 proteins were expressed in *Escherichia coli* and purified to homogeneity, as described previously<sup>3</sup>. Crystallization was performed at 25 °C, using the sitting-drop vapour-diffusion method. Crystals were obtained by mixing 1 µl of protein solution, consisting of 14 mg ml<sup>-1</sup> ST0318, 20 mM Tris-HCl (pH 7.5), 5 mM DHAP and 5 mM MgCl<sub>2</sub>, and 1 µl of reservoir solution, consisting of 0.1 M Bicine-KOH (pH 9.0), 10% PEG 20,000 and 2% dioxane. X-ray diffraction data were collected at the NW12A station ( $\lambda = 1.0 \text{ \AA}$ ) at the Photon Factory AR, High Energy Accelerator Research Organization, Tsukuba, Japan. The crystal structure of ST0318 in the FBPAse form (PDB code 1UMG) was used as the initial model for refinement. Data collection and refinement statistics are provided in Supplementary Table 1. Site-directed mutagenesis was performed with a PrimeSTAR mutagenesis kit (Takara Bio). The FBPA and FBPAse activities were both measured at 48 °C using a coupled spectrophotometric assay, as described previously<sup>2</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 25 March; accepted 15 August 2011.**

**Published online 9 October 2011.**

1. Moore, B. Bifunctional and moonlighting enzymes: lighting the way to regulatory control. *Trends Plant Sci.* **9**, 221–228 (2004).
2. Say, R. F. & Fuchs, G. Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* **464**, 1077–1081 (2010).
3. Nishimasu, H., Fushinobu, S., Shoun, H. & Wakagi, T. The first crystal structure of the novel class of fructose-1,6-bisphosphatase present in thermophilic archaea. *Structure* **12**, 949–959 (2004).
4. Rashid, N. *et al.* A novel candidate for the true fructose-1,6-bisphosphatase in archaea. *J. Biol. Chem.* **277**, 30649–30655 (2002).

5. Sato, T. *et al.* Genetic evidence identifying the true gluconeogenic fructose-1,6-bisphosphatase in *Thermococcus kodakaraensis* and other hyperthermophiles. *J. Bacteriol.* **186**, 5799–5807 (2004).
6. St-Jean, M. & Sygusch, J. Stereospecific proton transfer by a mobile catalyst in mammalian fructose-1,6-bisphosphate aldolase. *J. Biol. Chem.* **282**, 31028–31037 (2007).
7. St-Jean, M., Lafrance-Vanasse, J., Liotard, B. & Sygusch, J. High resolution reaction intermediates of rabbit muscle fructose-1,6-bisphosphate aldolase: substrate cleavage and induced fit. *J. Biol. Chem.* **280**, 27262–27270 (2005).
8. Choi, K. H., Shi, J., Hopkins, C. E., Tolan, D. R. & Allen, K. N. Snapshots of catalysis: the structure of fructose-1,6-(bis)phosphate aldolase covalently bound to the substrate dihydroxyacetone phosphate. *Biochemistry* **40**, 13868–13875 (2001).
9. Huber, R. & Bennett, W. S. Jr. Functional significance of flexibility in proteins. *Biopolymers* **22**, 261–279 (1983).
10. Bustamante, C., Chemla, Y. R., Forde, N. R. & Izhaky, D. Mechanical processes in biochemistry. *Annu. Rev. Biochem.* **73**, 705–748 (2004).
11. Tousignant, A. & Pelletier, J. N. Protein motions promote catalysis. *Chem. Biol.* **11**, 1037–1042 (2004).
12. Due, A. V., Kuper, J., Geerlof, A., Kries, J. P. & Wilmanns, M. Bisubstrate specificity in histidine/tryptophan biosynthesis isomerase from *Mycobacterium tuberculosis* by active site metamorphosis. *Proc. Natl Acad. Sci. USA* **108**, 3554–3559 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank O. Nureki for discussions, and the staff of the Photon Factory for the X-ray data collection.

**Author Contributions** S.F. collected the diffraction data and determined the crystal structure. H.N. prepared the wild-type and mutant proteins and crystallized the wild-type protein. D.H. prepared and crystallized the wild-type protein. D.H. and H.-J.S. measured the enzyme activity. S.F., H.N. and T.W. conceived the project and wrote the manuscript. All authors commented on the manuscript.

**Author Information** The atomic coordinates and structure factors are deposited in Protein Data Bank under accession number 3R1M. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to T.W. ([atwakag@mail.ecc.u-tokyo.ac.jp](mailto:atwakag@mail.ecc.u-tokyo.ac.jp)).

## METHODS

**Protein preparation and crystallography.** Wild-type and mutant ST0318 proteins were expressed in *E. coli* and purified to homogeneity, as described previously<sup>3</sup>. Crystallization was performed at 25 °C, using the sitting-drop vapour-diffusion method. Crystals were obtained by mixing 1 µl of protein solution, consisting of 14 mg ml<sup>-1</sup> ST0318, 20 mM Tris-HCl (pH 7.5), 5 mM DHAP and 5 mM MgCl<sub>2</sub>, and 1 µl of reservoir solution, consisting of 0.1 M Bicine-KOH (pH 9.0), 10% PEG 20,000 and 2% dioxane. X-ray diffraction data were collected at the NW12A station ( $\lambda = 1.0 \text{ \AA}$ ) at the Photon Factory AR, High Energy Accelerator Research Organization, Tsukuba, Japan. Crystals were cryoprotected in the reservoir solution supplemented with 25% 2-methyl-2,4-pentanediol, and were flash-cooled at 100 K in a stream of nitrogen gas. Data were processed using HKL2000 (ref. 13). The previously determined ST0318 structure (PDB code 1UMG) was used as the initial model for refinement. Manual model rebuilding and refinement were performed using Coot<sup>14</sup> and Refmac5 (ref. 15). The final model contains residues 2–353 and 361–364, 264 water molecules, one 2-methyl-2,4-pentanediol, three Mg<sup>2+</sup> ions and one DHAP molecule (O2 atom dehydrated) covalently attached to Lys 232. Data collection and refinement statistics are provided in Supplementary Table 1. Molecular graphic images were prepared using PyMol (Delano Scientific).

**Site-directed mutagenesis and enzyme assay.** Site-directed mutagenesis was performed with a PrimeSTAR mutagenesis kit (Takara Bio). The FBPA and FBPAse activities were both measured at 48 °C using a coupled spectrophotometric assay, as described<sup>2</sup>.

**FBPase assay.** FBP-dependent fructose-6-phosphate formation was measured by coupling the reaction with exogenous phosphoglucose isomerase and glucose-6-phosphate dehydrogenase, and NADPH formation was monitored at 340 nm ( $\epsilon_{340 \text{ nm}} \text{ NADPH} = 6,300 \text{ M}^{-1} \text{ cm}^{-1}$ ). The assay mixture (0.5 ml) consisted of 0.1 M Tris-HCl (pH 7.8), 20 mM MgCl<sub>2</sub>, 20 mM dithiothreitol (DTT), 0.5 mM NADP<sup>+</sup>, 0.01–0.3 mM FBP and 1 U each of phosphoglucose isomerase and glucose-6-phosphate dehydrogenase from baker's yeast (Sigma-Aldrich). The reaction was started by the addition of the purified enzyme.

**FBPA assay (anabolic direction).** Triosephosphate-dependent fructose-6-phosphate formation was measured by coupling the reaction with exogenous

triosephosphate isomerase, phosphoglucose isomerase and glucose-6-phosphate dehydrogenase from baker's yeast (Sigma-Aldrich); NADPH formation was monitored at 340 nm. The assay mixture (0.5 ml) consisted of 0.1 M Tricine-KOH (pH 8.0), 20 mM MgCl<sub>2</sub>, 20 mM DTT, 0.5 mM NADP<sup>+</sup>, 4 U of triosephosphate isomerase, 1 U of phosphoglucose isomerase and 1 U of glucose-6-phosphate dehydrogenase. The assay mixture was preincubated for 4 min. After the addition of GA3P (0.04–0.35 mM), the assay mixture was further incubated for 1 min to achieve equilibrium between GA3P and DHAP. The reaction was started by the addition of the purified enzyme.

**FBPA assay (catabolic direction).** FBP-dependent formation of triosephosphates was measured by coupling the reaction with triosephosphate isomerase from baker's yeast and glycerolphosphate dehydrogenase from rabbit (Sigma-Aldrich), and the oxidation of NADH was monitored at 365 nm ( $\epsilon_{365 \text{ nm}} \text{ NADH} = 3,400 \text{ M}^{-1} \text{ cm}^{-1}$ ). The assay mixture (0.5 ml) consisted of 0.1 M Tricine-KOH (pH 8.0), 20 mM MgCl<sub>2</sub>, 20 mM DTT, 0.55 mM NADH, 5 mM FBP, 20 U of triosephosphate isomerase and 2 U of glycerolphosphate dehydrogenase. The reaction was started by the addition of the purified enzyme.

To examine whether the enzyme is active under the crystallization conditions, we measured the FBPA activity (anabolic direction) in an assay mixture (0.5 ml) consisting of 0.1 M Bicine-KOH (pH 9.0), 10% PEG 20,000, 2% dioxane, 5 mM MgCl<sub>2</sub>, 0.5 mM NADP<sup>+</sup>, 0.5 mM GA3P, 0.5 mM DHAP, 20 U of triosephosphate isomerase, 5 U of phosphoglucose isomerase and 5 U of glucose-6-phosphate dehydrogenase. To examine whether the FBPA/P–DHAP complex represents a genuine reaction intermediate, we washed several crystals using the crystallization buffer, dissolved them in water and then measured the FBPA activity (anabolic direction).

- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).

# *Saccharomyces cerevisiae* THI4p is a suicide thiamine thiazole synthase

Abhishek Chatterjee<sup>1†</sup>, N. Dinuka Abeydeera<sup>2</sup>, Shridhar Bale<sup>1†</sup>, Pei-Jing Pai<sup>2</sup>, Pieter C. Dorrestein<sup>3</sup>, David H. Russell<sup>2</sup>, Steven E. Ealick<sup>1</sup> & Tadhg P. Begley<sup>2</sup>

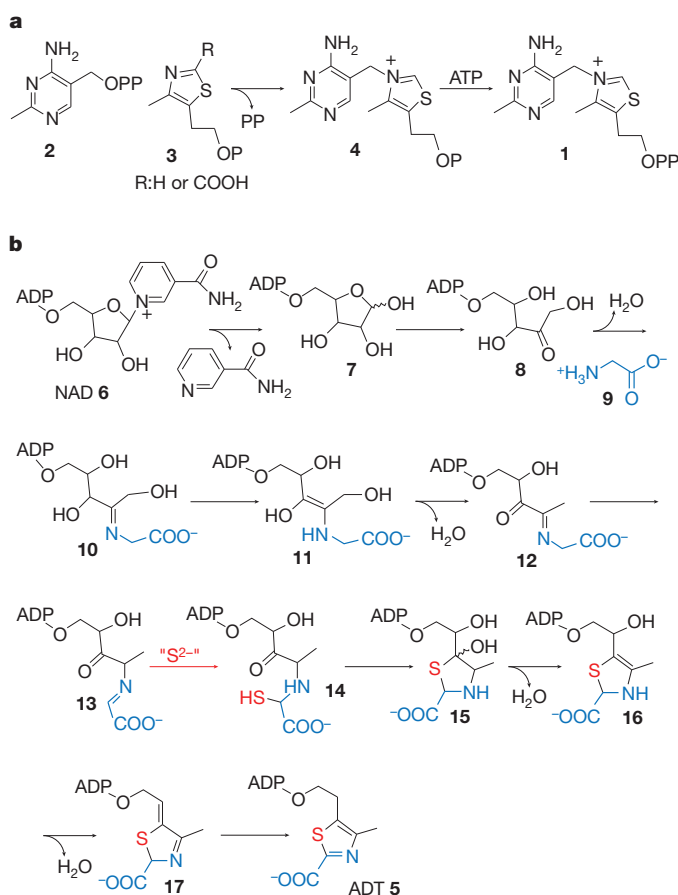
Thiamine pyrophosphate 1 is an essential cofactor in all living systems<sup>1</sup>. Its biosynthesis involves the separate syntheses of the pyrimidine 2 and thiazole 3 precursors, which are then coupled<sup>1</sup>. Two biosynthetic routes to the thiamine thiazole have been identified. In prokaryotes, five enzymes act on three substrates to produce the thiazole via a complex oxidative condensation reaction, the mechanistic details of which are now well established<sup>2-6</sup>. In contrast, only one gene product is involved in thiazole biosynthesis in eukaryotes (THI4p in *Saccharomyces cerevisiae*)<sup>7</sup>. Here we report the preparation of fully active recombinant wild-type THI4p, the identification of an iron-dependent sulphide transfer reaction from a conserved cysteine residue of the protein to a reaction intermediate and the demonstration that THI4p is a suicide enzyme undergoing only a single turnover.

While analysing the metabolic pathway leading to thiamine pyrophosphate **1** (Fig. 1a), we identified three adenylated metabolites (structures **5**, **12** and **17** in Fig. 1b), co-purifying with THI4p. They provide three molecular snapshots of the reaction pathway catalysed by this protein. In addition, two partially active mutants were identified (C204A and H200N), which catalysed the conversion of NAD (nicotinamide adenine dinucleotide) **6** and glycine **9** to an advanced intermediate **12** (ref. 8). A mechanism for thiazole formation, consistent with these observations, is outlined in Fig. 1b (refs 8–11). However, the source of the thiazole sulphur remained elusive, precluding us from deciphering the subsequent steps leading to the adenylated thiazole **5**.

High-resolution ESI-FTMS (electrospray ionization Fourier-transformed mass spectrometry) analysis of wild-type THI4p (wtTHI4p), recombinantly expressed in *Escherichia coli*, revealed a mass that was  $34 \pm 1$  Da lower than the calculated mass of the protein. The active site mutants of THI4p<sup>8</sup>, which did not copurify with any bound metabolites and did not show any activity, were unmodified, indicating that the 34-Da mass loss was in some way related to the catalytic activity of the protein. To localize the site of this modification, chymotrypsin digestion of modified wtTHI4p was carried out, followed by MALDI (matrix-assisted laser desorption ionization) and ESI-MS analysis of the peptide fragments. Before the digestion, free thiol residues of the protein were alkylated with iodoacetamide to protect them from oxidation. As a control, we performed the same procedure on an inactive and unmodified THI4p mutant (R301Q)<sup>8</sup> in parallel. Upon comparing the results for the wtTHI4p and the mutant THI4p, two modified peptide fragments, spanning the same region of the protein sequence, were identified (Fig. 2a and Supplementary Fig. 2). Fragmentation analysis localized the modification to a pair of adjacent cysteine residues (Cys 204 and Cys 205, Fig. 2a, highlighted in red). Both of these residues failed to alkylate during the iodoacetamide treatment of the peptide fragments. In contrast, peptide fragments originating from THI4p (R301Q) under the same conditions were completely alkylated (Fig. 2a). These observations may be explained by the transfer of H<sub>2</sub>S from Cys 204 or Cys 205 of wtTHI4p to a thiazole intermediate,

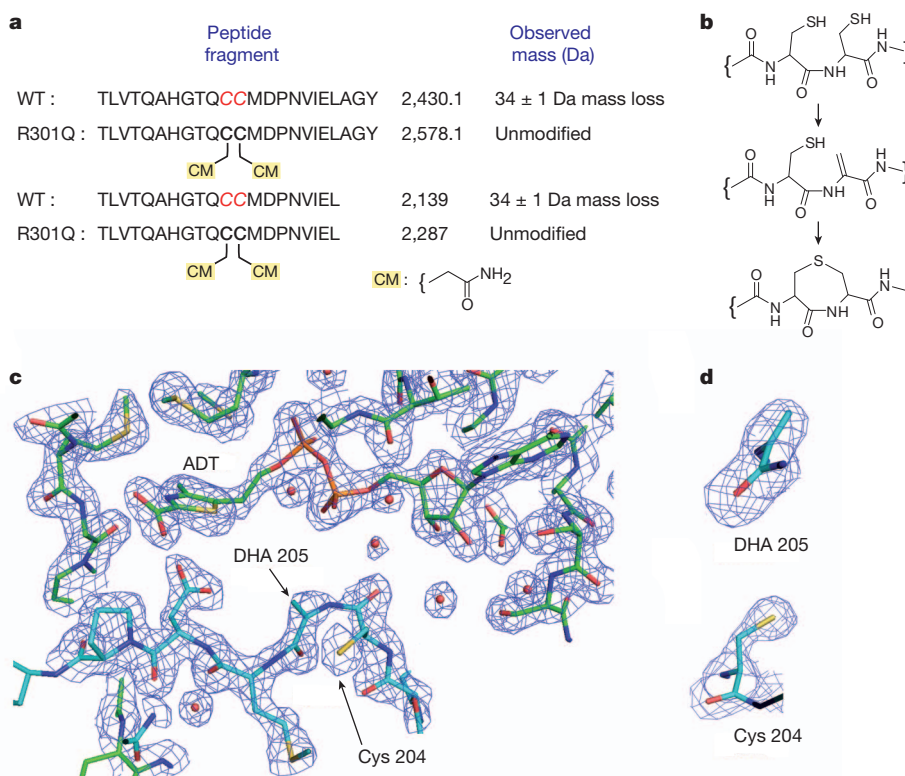
generating a dehydroalanine residue (M – 34 Da), which is subsequently trapped by the adjacent cysteine-thiol, producing a seven-membered cyclic thioether (Fig. 2b and Supplementary Fig. 2).

On the basis of these observations, the crystal structures of THI4p<sup>9</sup> and THI1 (ref. 12; THI4p orthologue from *Arabidopsis thaliana*) were reanalyzed for evidence of an active site dehydroalanine (DHA) or cyclic thioether. The  $2F_o - F_c$  and  $F_o - F_c$  electron density maps clearly demonstrate a lack of electron density for the sulphur atom of Cys 205, consistent with a dehydroalanine residue (Fig. 2c, d). The loop containing Ala 199–Asp 207, disordered in our previous structure, was completed and shown to extend into the active site of a fourfold-related monomer. The interpretation of the high resolution



**Figure 1 | Thiamine pyrophosphate and thiamine thiazole biosynthesis.**  
**a**, The late steps in thiamine pyrophosphate biosynthesis. **b**, Mechanistic proposal for the biosynthesis of thiamine-thiazole in eukaryotes catalysed by THI4p.

<sup>1</sup>Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA. <sup>2</sup>Department of Chemistry, Texas A&M University, College Station, Texas 77843, USA. <sup>3</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92093, USA. †Present address: The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037, USA.



**Figure 2 | Identification of the site of the M - 34 Da modification in wtTHI4p.** **a**, The peptide fragments originating from wtTHI4p and the R301Q mutant containing the site of modification. WT, wild type. **b**, A mechanistic hypothesis to explain the mass loss and lack of reactivity with iodoacetamide of the modified peptides. **c**, Active site of THI4p with bound ADT. The separation of the sulphur atom of ADT and the C $\beta$  atom of the dehydroalanine residue is

5.3 Å. The loop Gln 203–Pro 208 is from a fourfold-related monomer and has carbon atoms coloured cyan. Water molecules are shown as red spheres. The electron density map ( $2F_o - F_c$  contoured at  $3\sigma$ ) clearly shows the loss of sulphur from Cys 205 to form the dehydroalanine residue. **d**, Magnified electron density of residue DHA 205 and residue Cys 204.

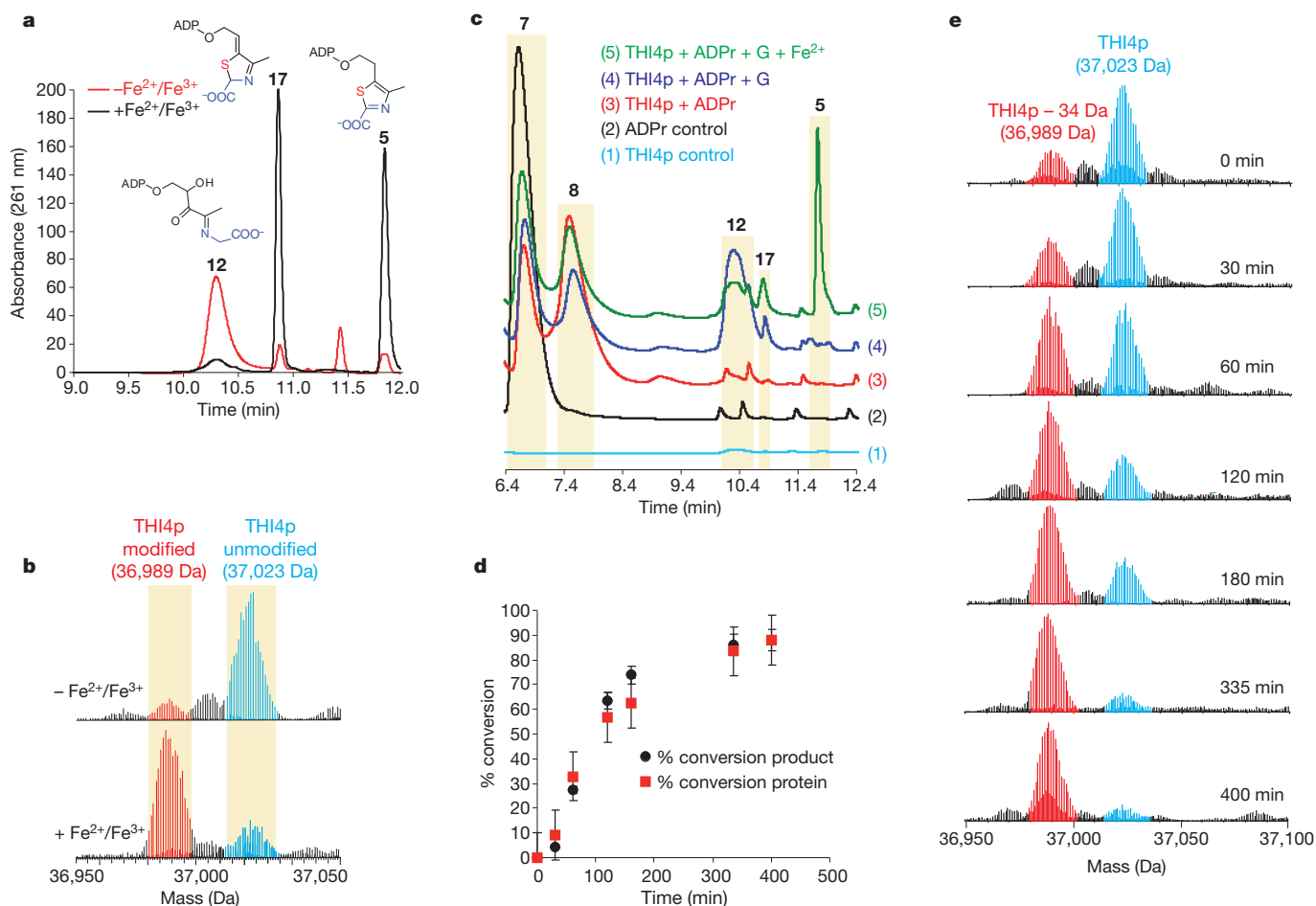
structure of THI1 (PDB ID 1RP0) is also consistent with an active site dehydroalanine. The  $2F_o - F_c$  and  $F_o - F_c$  electron density maps for this protein confirm that electron density for the sulphur atom of Cys 172 is missing and demonstrate that the C $\alpha$  is planar (Supplementary Fig. 3). The identification of adjacent cysteine/dehydroalanine residues in THI4p suggests that the formation of the cyclic thioether occurs under the denaturing conditions used for the preparation of the sample for mass spectrometry analysis.

The activity-dependent loss of H $_2$ S from THI4p indicated that the sulphur atom of the thiazole could be derived from THI4p. The resulting modification renders the enzyme inactive and explains our previous inability to reconstitute active wtTHI4p. The problem of obtaining active wtTHI4p was solved by a surprising set of observations regarding the effect of the growth medium on the activity of isolated wtTHI4p. When the overexpression strain was grown in M9 minimal medium instead of Luria–Bertani (LB) medium, the purified protein was mostly free of the 34-Da mass modification (Fig. 3b). High-performance liquid chromatography (HPLC) analysis of this protein preparation demonstrated greatly reduced formation of the sulphur-containing metabolites **17** and **5** and the accumulation of **12** (Fig. 3a, red HPLC trace). This indicates that the sulphur transfer chemistry involved in thiazole formation is greatly retarded when wtTHI4p is isolated from cells grown in minimal medium.

Addition of 100  $\mu$ M iron (II) to the growth medium was sufficient to restore the sulphide transfer activity of wtTHI4p as indicated by the formation of **17** and **5** shown in Fig. 3a (black HPLC trace). In addition, the protein thus isolated showed the 34-Da mass loss, as observed with wtTHI4p overexpressed in LB medium (Fig. 3b). These observations indicated that the THI4p-catalysed sulphur incorporation reactions are iron-dependent.

These observations suggested conditions for the successful reconstitution of the THI4p-catalysed reaction. To accomplish this, the unmodified wtTHI4p, overexpressed in minimal media, was freed from bound metabolites by multiple rounds of gel-filtration. This protein preparation catalysed the conversion of ADP-ribose (ADPr **7**) to **12** in the presence of glycine **9**, via the intermediate of ADP-ribulose (ADPrI **8**, Fig. 3c). Addition of Fe $^{+2}$  to this reaction mixture resulted in the conversion of **12** to the final product ADP-thiazole (ADT, **5**; Fig. 3c). Other divalent metal ions (Mg $^{+2}$ , Ca $^{+2}$ , Mn $^{+2}$ , Co $^{+2}$ , Ni $^{+2}$ , Cu $^{+2}$  and Zn $^{+2}$ ) did not activate the enzyme for sulphide transfer chemistry. Although this reaction mixture did not contain an exogenous sulphur source, nearly one full turnover could be observed ( $380 \pm 5 \mu$ M ADT from  $420 \pm 21 \mu$ M THI4p). Inclusion of excess sulphide or cysteine in the reaction mixture did not enhance the turnover number or protect the protein from modification. The production of ADT **5** was oxygen-sensitive and had to be performed under an anaerobic atmosphere, in the presence of the reducing agent dithiothreitol (DTT) or tris-(2-carboxyethyl)-phosphine (TCEP). Formation of ADT was accompanied by the loss of H $_2$ S ( $\Delta M = 34$  Da) from wtTHI4p, as evidenced from ESI-FTMS analysis (Fig. 3e). In control reactions, lacking either Fe $^{+2}$  or ADPr **7**, no modification of wtTHI4p was observed. A time course for the reaction demonstrates a consistent stoichiometry of 1:1 for protein modification and thiazole production (Fig. 3d, e).

The full reconstitution of ADT **5** formation, using purified wtTHI4p, in a reaction mixture lacking any exogenous sulphide donor, coupled with the observed loss of 34 Da from the protein and the structural characterization of a dehydroalanine residue at Cys 205 provides compelling evidence that the thiazole sulphur is derived from Cys 205 of wtTHI4p. Consistent with this model is the observation that Cys 205 is strictly conserved in eukaryotic thiazole synthases, whereas



**Figure 3 | Reconstitution of the biosynthesis of ADT 5.** **a**, HPLC analysis of the metabolites associated with wtTHI4p overexpressed in M9 minimal medium  $\pm$  100  $\mu$ M iron. **b**, ESI-FTMS analysis of wtTHI4p overexpressed in M9 minimal medium  $\pm$  100  $\mu$ M iron shows iron-dependent modification ( $\Delta M = -34$  Da) of the protein. **c**, HPLC analysis of wtTHI4p-catalysed partial and full reactions and the relevant control reactions. Incubating THI4p with

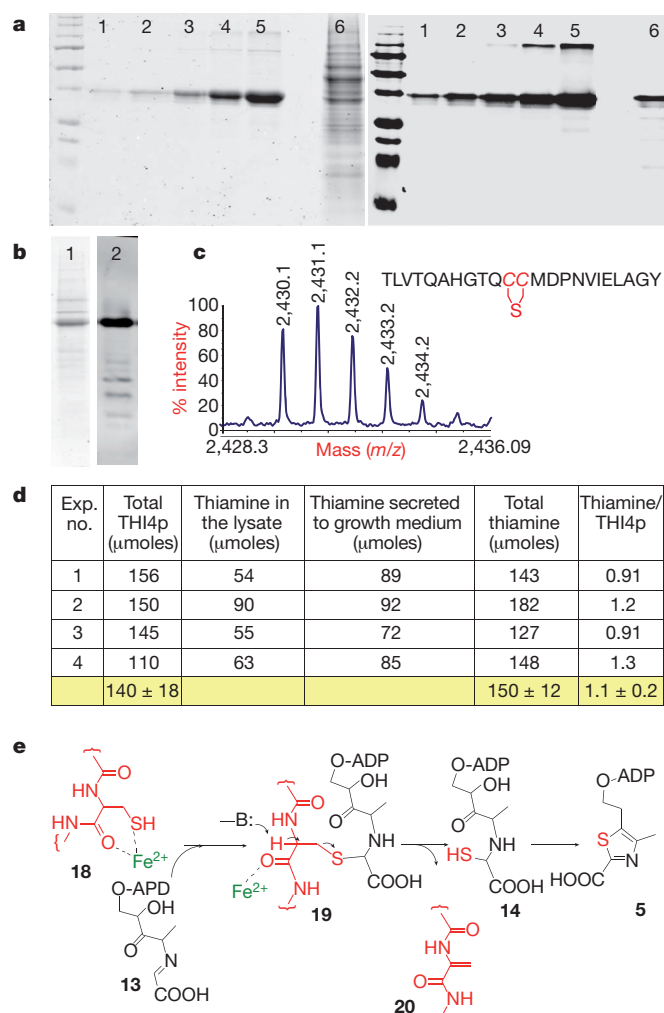
ADPr 7, glycine and iron (II) results in the production of ADT 5. **d**, Time course for the reaction showing a 1:1 ratio of protein modification and thiazole production (error bars indicate s.d.). **e**, Mass spectrometry analysis of wtTHI4p over the time course of the reaction showing the progressive conversion of the enzyme to the M - 34 Da species.

Cys 204 is often replaced with a serine residue. Also, the THI4p C205S mutant was shown to be inactive, whereas the C204S mutant retains its ability to produce ADT, indicating Cys 205 as the sulphur donating residue.

Efforts towards the *in vitro* reactivation of modified THI4p under a variety of different conditions were unsuccessful. This suggests that THI4p is a single turnover enzyme. In addition, whereas thiamine biosynthetic enzymes are generally present in the proteome at very low concentrations, THI4p is an exception and a high level of expression (approximately 1.5%)<sup>13</sup> is observed during the exponential growth phase of *Neurospora crassa*. To investigate the possibility that THI4p is a single turnover enzyme, we first characterized THI4p, expressed at native levels in a yeast strain, in which a His<sub>6</sub> tag had been inserted at the carboxy terminus of THI4p to facilitate its detection and purification. THI4p was isolated from this strain grown to mid-log phase in a vitamin-free defined medium using immunoaffinity chromatography and was analysed by gel electrophoresis and western blotting (Fig. 4a, b). The THI4p band was excised, digested with chymotrypsin and the peptide fragments were subjected to mass spectrometric analysis (MALDI-TOF/TOF MS). The peptide fragment (T194 to Y216) associated with the 34 Da mass-loss ( $m/z = 2,430.1$ ) was again identified in this experiment (Fig. 4c) and their fragmentation patterns were identical, confirming the loss of H<sub>2</sub>S followed by cyclic thioether formation (Supplementary Figs 4 and 5). The unmodified species for the same peptide (predicted  $m/z = 2,464$  Da) was not observed in this

experiment. Identification of the same pattern of H<sub>2</sub>S-loss from native THI4p, as observed in the *in vitro* reconstitution experiments, further validates THI4p as the sulphur donor for thiazole biosynthesis *in vivo*. In addition, the mass spectrometric analysis identifies modified THI4p as the major species present *in vivo*, indicating the absence of a repair pathway in yeast for modified THI4p. This suggests that the ratio of THI4p to thiamine produced should be 1:1. To test this hypothesis, the concentration of native THI4p in yeast cell-free extract was determined by quantitative western blot analysis, using known quantities of recombinantly expressed THI4p to generate a calibration curve (Fig. 4a), and thiamine concentration in yeast lysate and in the growth medium (secreted by yeast) was determined by the thiochrome assay (supplementary Fig. 6). A  $1:1.1 \pm 0.2$  stoichiometry between THI4p and thiamine was demonstrated in yeast cultures growing at mid-log phase (four independent experiments), supporting its role as a single turn-over enzyme (Fig. 4d).

A mechanistic proposal for the sulphide transfer chemistry involved in ADT 5 formation that is consistent with these observations is outlined in Fig. 4e. Addition of the thiol of Cys 205 to intermediate 13, formed as shown in Fig. 1, followed by  $Fe^{+2}$  assisted elimination would give 14 and generate the active site dehydroalanine observed in the structure. It is also possible that the iron activates the sulphide transfer by direct interaction with the sulphur. The oxygen sensitivity of the reaction indicates that  $Fe^{+2}$  is the catalytically active iron oxidation state. Intermediate 14 is then converted to ADT 5 as shown in Fig. 1.



**Figure 4 | Characterization of native THI4p from *S. cerevisiae*.** **a**, Analysis of THI4p in yeast cell-free extract using a western blot (right) and Coomassie blue (left). Lanes 1–5 contain increasing concentrations (0.5, 1, 2, 4 and 8 μM) of His<sub>6</sub>-THI4p. Lane 6 contains yeast crude lysate. **b**, Native THI4p, isolated from yeast cell-free extract, analysed by SDS-PAGE/Coomassie blue (1) and western blot (2). **c**, In-gel chymotrypsin digestion/MALDI-TOF analysis of isolated native THI4p demonstrates that the peptide containing the C204–C205 region has the same modification as observed with THI4p expressed in *E. coli*. **d**, Quantification of THI4p and thiamine produced in a culture of yeast, grown in vitamin-free defined medium, demonstrates that THI4p is a substrate rather than a catalyst. **e**, Proposed mechanism for the iron-mediated sulphur transfer reaction involved in the formation of intermediate 14 (Fig. 1).

The observations reported here strongly indicate that THI4p acts as a co-substrate rather than an enzyme. This is very unusual but is not without precedent. The best characterized example of a single turnover enzyme is the Ada protein which repairs O<sup>6</sup>-methylguanine and methyl-phosphotriester lesions in DNA by transferring the methyl group to an active site cysteine. The resulting inactive enzyme serves as a signal to induce other DNA repair enzymes<sup>14–16</sup>. The possibility that inactive THI4p has other physiological function(s) remains to be explored. Interestingly, involvement of THI4p and its orthologues has been implicated in DNA protection and other stress-related pathways<sup>17–19</sup>. The mechanism of this protection is not known. One possibility is that the abundant THI4p protects the cells by binding free cellular iron, which is known to cause oxidative damage via the generation of reactive oxygen species.

THI4p is a remarkable protein: it is a suicide enzyme serving as a co-substrate rather than an enzyme for the formation of the thiazole moiety of thiamine. This assembly involves a complex, unprecedented reaction sequence in which NAD serves as the source of the five-carbon

chain and THI4p serves as the sulphur source. The biological function of the modified protein, if any, remains to be elucidated.

## METHODS SUMMARY

THI4p and its mutants (pET28b vector, BL21(DE3) cell line) were overexpressed in LB medium and isolated using Ni-NTA chromatography as described previously<sup>8</sup>. To isolate unmodified THI4p for activity assays, M9 minimal medium was used for overexpression (see Supplementary Materials for details). HPLC analyses of THI4p-bound metabolites and the *in vitro* reconstitution assays were performed as previously described<sup>8</sup>. Assays of THI4p were performed in an anaerobic chamber (COY Laboratory Products Inc.). A typical THI4p reconstitution reaction assay included ADP-ribose or NAD (final concentration 1 mM), glycine (final concentration 1 mM), freshly prepared FeSO<sub>4</sub> (final concentration 0.5 mM) and THI4p (300–500 μM). THI4p and its mutants were desalted into 50% methanol containing 0.1% formic acid for ESI-FTMS analyses. Chymotrypsin digestion/MALDI analyses of recombinant THI4p and its mutants were performed at the Proteomics Facility, Cornell University, using standard protocols. In-gel chymotrypsin digestion/MALDI-TOF/TOF MS analysis with THI4p isolated from yeast was carried out in the Russell laboratory at Texas A&M University and the ESI-FTMS analysis was carried out in the Dorrestein laboratory at UCSD. The crystal structures of THI4p (Protein Data Bank ID 2GJC)<sup>9</sup> and THI1p (PDB ID 1RP0)<sup>12</sup> were refined against the deposited structure factor magnitudes using REFMAC from CCP4 (refs 20 and 21). The *Saccharomyces cerevisiae* strain BY4741 (purchased from Open Biosystems, Inc.) was used for *in vivo* studies with THI4p. Vitamin-free medium A (medium YNB (yeast nitrogen base) plus 0.01% uracil, 0.01% leucine, 0.005% histidine, 0.005% methionine, 2% glucose, and 0.03% G418) was used to grow the yeast strain to allow the expression of the endogenous THI4p. Endogenous THI4p concentration was determined using a quantitative western blot, whereas the concentration of thiamine was determined using the thiochrome assay<sup>10</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 4 May; accepted 24 August 2011.

- Jordan, F. in *Comprehensive Natural Products Chemistry II: Chemistry and Biology* (eds Mander, L. & Liu, H.-W.) 561–598 (Elsevier, 2010).
- Begley, T. P. Cofactor biosynthesis: an organic chemist's treasure trove. *Nat. Prod. Rep.* **23**, 15–25 (2006).
- Begley, T. P. & Ealick, S. E. in *Comprehensive Natural Products Chemistry II: Chemistry and Biology* (eds Mander, L. & Liu, H.-W.) 547–560 (Elsevier, 2010).
- Jurgenson, C. T., Begley, T. P. & Ealick, S. E. The structural and biochemical foundations of thiamin biosynthesis. *Annu. Rev. Biochem.* **78**, 569–603 (2009).
- Hazra, A. B. *et al.* Missing enzyme in thiamin thiazole biosynthesis: identification of Ten1 as a thiazole tautomerase. *J. Am. Chem. Soc.* **133**, 9311–9319 (2011).
- Kriek, M. *et al.* Thiazole synthase from *Escherichia coli*: an investigation of the substrates and purified proteins required for activity *in vitro*. *J. Biol. Chem.* **282**, 17413–17423 (2007).
- Prækel, U. M., Byrne, K. L. & Meacock, P. A. Regulation of *THI4* (MOL1), a thiamine-biosynthetic gene of *Saccharomyces cerevisiae*. *Yeast* **10**, 481–490 (1994).
- Chatterjee, A., Jurgenson, C. T., Schroeder, F. C., Ealick, S. E. & Begley, T. P. Biosynthesis of thiamin thiazole in eukaryotes: conversion of NAD to an advanced intermediate. *J. Am. Chem. Soc.* **129**, 2914–2922 (2007).
- Jurgenson, C. T., Chatterjee, A., Begley, T. P. & Ealick, S. E. Structural insights into the function of the thiamin biosynthetic enzyme Thi4 from *Saccharomyces cerevisiae*. *Biochemistry* **45**, 11061–11070 (2006).
- Chatterjee, A., Jurgenson, C. T., Schroeder, F. C., Ealick, S. E. & Begley, T. P. Thiamin biosynthesis in eukaryotes: characterization of the enzyme-bound product of thiazole synthase from *Saccharomyces cerevisiae* and its implications in thiazole biosynthesis. *J. Am. Chem. Soc.* **128**, 7158–7159 (2006).
- Chatterjee, A., Schroeder, F. C., Jurgenson, C. T., Ealick, S. E. & Begley, T. P. Biosynthesis of the thiamin-thiazole in eukaryotes: identification of a thiazole tautomer intermediate. *J. Am. Chem. Soc.* **130**, 11394–11398 (2008).
- Godoi, P. H. *et al.* Structure of the thiazole biosynthetic enzyme THI1 from *Arabidopsis thaliana*. *J. Biol. Chem.* **281**, 30957–30966 (2006).
- Faou, P. & Tropschug, M. *Neurospora crassa* CyBP37: a cytosolic stress protein that is able to replace yeast THI4p function in the synthesis of vitamin B1. *J. Mol. Biol.* **344**, 1147–1157 (2004).
- Teo, I., Sedgwick, B., Kilpatrick, M. W., McCarthy, T. V. & Lindahl, T. The intracellular signal for induction of resistance to alkylating agents in *E. coli*. *Cell* **45**, 315–324 (1986).
- Sedgwick, B., Robins, P., Totty, N. & Lindahl, T. Functional domains and methyl acceptor sites of the *Escherichia coli* Ada protein. *J. Biol. Chem.* **263**, 4430–4433 (1988).
- Demple, B. *et al.* Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis. *Proc. Natl Acad. Sci. USA* **82**, 2688–2692 (1985).
- Machado, C. R. *et al.* Dual role for the yeast *THI4* gene in thiamine biosynthesis and DNA damage tolerance. *J. Mol. Biol.* **273**, 114–121 (1997).

18. Ruiz-Roldán, C. *et al.* The *Fusarium oxysporum sti35* gene functions in thiamine biosynthesis and oxidative stress response. *Fungal Genet. Biol.* **45**, 6–16 (2008).
19. Medina-Silva, R. *et al.* Heat stress promotes mitochondrial instability and oxidative responses in yeast deficient in thiazole biosynthesis. *Res. Microbiol.* **157**, 275–281 (2006).
20. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
21. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D* **55**, 247–255 (1999).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank F. W. McLafferty for initial FTMS analysis of THI4p. This research was funded by NIH grant DK44083 (to T.P.B.), the Robert E. Welch Foundation grant A-0034 (to T.P.B.), DK67081 (to S.E.E.) and NSF grant DBI0821700 (to D.H.R.).

**Author Contributions** A.C. performed all biochemical experiments with recombinantly expressed THI4p, N.D.A. performed all biochemical experiments with endogenous THI4p from yeast, S.B. performed structural analyses, P.C.D. supervised FTMS experiments and assisted in FTMS data analyses, and P.-J.P. performed the sequence analysis by mass spectrometry on endogenous THI4p. D.H.R. supervised the mass spectrometric analysis on endogenous THI4p, S.E.E. supervised the structural studies and T.P.B. supervised the biochemical studies.

**Author Information** Atomic coordinates for the crystal structure of THI4p reported here have been deposited in the Protein Data Bank under accession code 3FPZ. This entry includes structure factors originally contained within PDB accession code 2GJC. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to T.P.B. ([begley@chem.tamu.edu](mailto:begley@chem.tamu.edu)) or to S.E.E. ([see3@cornell.edu](mailto:see3@cornell.edu)).

## METHODS

All the chemicals were purchased from Sigma-Aldrich Corporation (USA) unless otherwise mentioned. LB medium was obtained from EMD Biosciences. Yeast nitrogen base (YNB) with amino acids and ammonium sulphate, vitamin-free, without dextrose (powder) was purchased from US Biological, Inc. Kanamycin and isopropyl- $\beta$ -D-thiogalactoside (IPTG) were purchased from Lab Scientific Inc. Geneticin (G418 sulphate) was obtained from GIBCO-Invitrogen. Zymolyase-100T and chymotrypsin were obtained from AMS Biotechnology (Europe) Ltd and Promega, respectively. The Penta-His Alexa Fluor 647 antibody conjugate was purchased from Qiagen. The alkaline phosphatase, calf intestinal (CIP) was obtained from New England Biolabs Inc. Analytical HPLC was carried out using a Supelco LC-18-T (150  $\times$  4.6 mm internal diameter; particle size 3  $\mu$ m) reverse-phase column. HPLC-grade solvents were obtained from Fisher Scientific. HPLC analyses were performed in an Agilent 1200 instrument equipped with an inline diode-array detector and a fluorescence detector.

**Protein expression and purification.** M9 minimal medium was prepared by autoclaving a 2 l solution of M9 minimal salts (22.5 g; Sigma-Aldrich) and upon cooling, complementing it with 25 ml of 50% sterilized glucose, 4 ml of 1 M sterilized  $\text{MgSO}_4$ , 0.2 ml of sterilized 1 M  $\text{CaCl}_2$  and 40  $\text{mg l}^{-1}$  kanamycin. A 10-ml culture of BL21(DE3) cells containing the plasmid used to express wtTHI4p was used to inoculate 2 l M9 minimal medium prepared as described above. Cultures were agitated at 37 °C until the OD at 600 nm reached 0.6, at which point protein expression was induced by adding 1 mM IPTG. The temperature was reduced to 15 °C and the culture was grown overnight with shaking. The cultures were then harvested by centrifugation (14,000g) and the pellets were frozen and stored at -80 °C until further use. To demonstrate iron dependence, an identical growth protocol was used except that, immediately before the induction of protein expression,  $\text{FeSO}_4$  was added to the culture (final concentration 100  $\mu\text{M}$ ).

To purify the protein, a cell pellet from 2 l of culture was resuspended in 30 ml lysis buffer (10 mM imidazole, 300 mM NaCl, 50 mM  $\text{NaH}_2\text{PO}_4$ , 1 mM DTT, pH 8) and lysed by sonication on ice (Heat Systems Ultrasonics model W-385 sonicator, 2-s cycle, 50% duty). The resulting cell lysate was clarified by centrifugation (30,000g) and the THI4p protein was purified using a 5 ml HisTrap column (GE Healthcare) following the manufacturer's instructions. After elution, the protein was buffer-exchanged using a 10-DG column (BioRad) pre-equilibrated with 50 mM potassium phosphate buffer, pH 7.8, containing 2 mM DTT.

To release THI4p-bound metabolites, this protein preparation was further gel-filtered into 20 mM Tris-HCl pH 7.8 (containing 2 mM DTT). Then it was buffer exchanged into 50 mM KPi, pH 7.8 containing 2 mM DTT (or TCEP) and 100 mM NaCl. This protein preparation was used for all assays.

**ESI-FTMS analysis of THI4p and its mutants.** Samples were prepared by desalting the protein using C18 ZipTip (Millipore) following the manufacturer's protocol and eluting with 75% acetonitrile containing 0.1% formic acid. This was diluted 1:1 with 50% methanol containing 0.1% formic acid and the resulting solution was used for ESI-FTMS analysis (Thermo-Fisher LTQ-FT) using a Biversa nanospray robot (Advion). The resulting data were deconvoluted with the software package Extract (licensed by Thermo-Fisher).

**Chymotrypsin digestion-MALDI analysis of recombinantly expressed, wtTHI4p and the R301Q mutant.** Guanidinium hydrochloride (40  $\mu\text{l}$ , 8 M solution in 25 mM Tris-HCl, pH 7.5) was added to a protein sample (10  $\mu\text{l}$ , 10  $\text{mg ml}^{-1}$ ), followed by 2.5  $\mu\text{l}$  of freshly prepared 200 mM DTT. After incubating the mixture for an hour at 50 °C, 19.5  $\mu\text{l}$  of freshly prepared 200 mM iodoacetamide solution was added to a final concentration of 55 mM and the reaction mixture was incubated for 1 h at room temperature in the dark. An aliquot of this reaction mixture (36  $\mu\text{l}$ ) was added to a solution of 2.5  $\mu\text{g}$  chymotrypsin in 150  $\mu\text{l}$  100 mM Tris-HCl containing 10 mM  $\text{CaCl}_2$ . The proteolysis reaction was allowed to proceed overnight at room temperature in the dark. The resulting peptides were desalted using a PrepSep C18 solid phase extraction cartridge (Fisher Scientific) according to the manufacturer's protocol and were subjected to MALDI-TOF/TOF MS analysis using a 4700 Proteomics Analyzer.

**In vitro reconstitution of wtTHI4p-catalysed reactions.** The wtTHI4p, over-expressed in minimal media, was transferred to an anaerobic chamber (Coy Laboratory Products, Inc.). ADP-ribose or NAD (final concentration 1 mM), glycine (final concentration 1 mM) and freshly prepared  $\text{FeSO}_4$  (final concentration 0.5 mM) were added to the protein (420  $\mu\text{M}$ ) to initiate the reaction. Control reactions were set up by omitting various essential components. Reactions were incubated at room temperature for 6 h, heat-quenched and analysed by HPLC as previously described<sup>8</sup>. The concentration of ADT produced was measured by converting it to the fluorescent thiochrome phosphate, as previously described<sup>10</sup>.

To observe the time-dependence of the THI4p-mediated reaction, a large scale reaction was set up (3 ml). Aliquots (100  $\mu\text{l}$ ) were heat-quenched at appropriate

time intervals and the quenched reaction mixture was analysed by HPLC as previously described<sup>8</sup>. Simultaneously, another 100  $\mu\text{l}$  aliquot was removed from the anaerobic chamber, rapidly buffer-exchanged, using a Bio-Spin mini desalting column (Bio-Rad), into 25 mM Tris-HCl, pH 7.5, frozen in liquid nitrogen and stored at -80 °C. Later, these samples were individually thawed, immediately before use, and desalted and eluted with 75% acetonitrile containing 0.1% formic acid for ESI-FTMS analysis (as described above).

**Identification of dehydroalanine in the structures of THI4p and THI1p.** Consistent with the biochemical studies, electron density maps showed no density for the sulphur atom of Cys 205 and this residue was changed to dehydroalanine. In addition, several surrounding residues were adjusted to fit the electron density and a loop missing from the original structure was built. New loops including Gln 198-Thr 202, His 237-Gly 242, Pro 178-Thr 185 and Asp 238-Phe 24 and loop Asp 219-His 228 were rebuilt to better fit the electron density. Model building was done using program COOT<sup>22</sup>. Analysis of the Ramachandran plot for THI4p showed that 91.8% of the non-glycine residues were in the most favoured region, 7.8% in the additional favoured region, 0.2% in the generously allowed region, and 0.2% in the disallowed region. The final refinement statistics are given in Supplementary Table 1.

The crystal structure of THI1p (PDB ID 1RP0)<sup>12</sup> was refined against the deposited structure factor magnitudes using REFMAC. The structure of THI1p was originally deposited with an annotation that the occupancy of the sulphur atom of Cys 172 was 0, indicating that electron density for the sulphur atom was not visible. Residue Cys 172 was changed to dehydroalanine and the structure was refined a few rounds to convergence. Analysis of the Ramachandran plot for THI1p showed that 92.9% of the non-glycine residues were in the most favoured region and the remaining 7.1% were in the additional favoured region. No residues were found in the generously allowed or disallowed regions. The final refinement statistics, which are very similar to those originally reported<sup>12</sup>, are given in Supplementary Table 1. Figure 2c, d was made using PyMOL<sup>23</sup>.

**Yeast strain and cell growth.** The *Saccharomyces cerevisiae* strain BY4741 (*MATa ura3 $\Delta$ 0 leu2 $\Delta$ 0 his3 $\Delta$ 1 met15 $\Delta$ 0*) was purchased from Open Biosystems, Inc. The *MATa* cells of the background strain BY4741 are tagged at the 3' end of the *THI4* ORF with a sequence encoding six histidines (6  $\times$  His)<sup>24</sup>. The chromosomal insertion was verified by colony PCR analysis and G418 resistance.

For each experiment, a single colony was picked from a freshly streaked G418 plate that was incubated at 30 °C for 2 days. For cell growth, 10 ml of YPD medium (1% yeast extract, 2% bacto peptone, 2% glucose) was inoculated with this colony and incubated in a 100 ml sterile Erlenmeyer flask at 30 °C (shaking speed 270 r.p.m.) for 12 h or until an attenuation of 4.0 at 600 nm was reached. The resulting cells were collected by centrifugation, washed (3 $\times$ ) with vitamin-free medium A (medium YNB (yeast nitrogen base) plus 0.01% uracil, 0.01% leucine, 0.005% histidine, 0.005% methionine, 2% glucose and 0.03% G418), suspended in 50 ml of vitamin-free medium A containing the antibiotic G418 to an attenuation of 0.07 at 600 nm and transferred to a 250 ml sterile Erlenmeyer flask. This was incubated for 15 h (the predicted mid-log phase) at 30 °C (shaking speed 270 r.p.m.) and the cells were collected by centrifugation at 4,000g for 10 min.

For experiments to determine the THI4p or thiamine concentration, cells (30–40 mg wet weight) were collected from 50 ml of vitamin-free medium A.

For the mass spectrometric experiments, the cells (5–10 g, wet weight) were collected from 6 l (3  $\times$  2 l) of vitamin-free medium A.

**Western blot analysis.** Yeast cell-free extract was analysed by SDS-PAGE, transferred to polyvinylidene difluoride (PVDF) membranes using a Panther semi-dry electroblotter (Owl Separation Systems), and then blocked in TBS (Tris-buffered saline, 50 mM Tris-HCl, pH 7.4 and 150 mM NaCl) plus 3% BSA and 0.05% Tween 20 for 1 h, before incubating with Penta-His Alexa Fluor 647 (Qiagen) anti-His antibody conjugates (1:5,000 dilution) for 1 h in the dark. After two rounds of 10-min washes with TBS in the dark, the membrane was then subjected to direct immunofluorescent detection by Typhoon Trio (excitation, 633 nm red laser; emission, 670 nm band-pass filter (670 BP 30)) from GE Healthcare Biosciences.

The above protocol was followed to confirm the occurrence and/or purity of endogenous THI4p isolated from yeast.

**Analysis of the THI4p/thiamine ratio in vivo.** The yeast cultures were grown as described above. The cells grown to the mid-log phase in 50 ml of vitamin-free medium A were harvested by centrifugation at 4,000g for 10 min. A portion (20 ml) of the supernatant (the clarified culture medium) was collected and frozen at -80 °C for at least 12 h. This was then lyophilized to dryness and the residue was used to determine the amount of thiamine secreted into the growth medium (details will follow).

The wet weight of cells from the 50 ml culture was 30 mg. Spheroplast formation was carried out according to published protocols<sup>25</sup>. Unless indicated otherwise, all steps were performed at 4 °C. Cells (30 mg) were re-suspended in 30  $\mu\text{l}$  of

zymolyase buffer 1 (50 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 30 mM TCEP), incubated for 15 min at room temperature, centrifuged at 400g for 5 min and the supernatant was discarded (thiamine concentration in this fraction was negligible). The cell pellet was re-suspended in 100 µl of zymolyase buffer 2 (50 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mM TCEP) and the cell wall was enzymatically digested with 2 mg ml<sup>-1</sup> of zymolyase-100T ( $1.0 \times 10^5$  U g<sup>-1</sup>), which contained β-1,3-glucan laminarinpentaohydrolase (from *Arthrobacter* sp.), β-1,3-glucanase (from *Arthrobacter* sp.), mannanase and protease. The suspension was incubated for 30 min at 30 °C with shaking at 60 r.p.m. Light microscopy confirmed complete spheroplast formation. The spheroplasts were then centrifuged at 400g for 5 min and the supernatant was collected in a 1.5 ml Eppendorf tube (Tube A). The spheroplasts were washed with 100 µl of ice-cold zymolyase buffer 2 followed by centrifugation at 400g for 5 min. and the supernatant was added to Tube A. The spheroplast pellet was resuspended in 100 µl of ice-cold lysis buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mM TCEP, 10 mM KOAc, 1 mM PMSF plus protease inhibitor cocktail (100 mM AEBSF, 500 mM 1,10-phenanthroline, 2.2 mM pepstatin A, and 1.4 mM E-64)). The suspension was homogenized on ice in a Dounce homogenizer (1 ml) by 20 strokes with a large clearance pestle "A" (Kontes Glassware). The spheroplast lysate was transferred to Tube A. An equal volume (300 µl) of extraction buffer (lysis buffer plus 0.8 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 20% glycerol) was added to Tube A. The tube was agitated on a rocker for 30 min and centrifuged for 30 min at 30,000g. The supernatant (600 µl) was then split into two equal portions. One portion (300 µl) was analysed for THI4p while the other was analysed for thiamine.

To analyse for THI4p, one of the above portions was concentrated to 100 µl and 5 µl from that concentrate was loaded onto a 12% SDS-PAGE along with recombinant THI4p standards (0.5, 1, 2, 4 and 8 µM). A standard curve was determined with each western blot. The SDS PAGE and quantitative western blot analysis were performed as described above. The THI4p band densities were analysed using the ImageQuant image analysis software provided with the Typhoon Trio.

The second 300 µl portion of the cell lysate was treated with 5 µl of calf intestinal alkaline phosphatase (CIP) and incubated at 37 °C for 18 h to convert the thiamine vitamers to thiamine alcohol. Potassium acetate (50 µl of 4 M) was added to 100 µl of the alkaline-phosphatase-treated sample followed by oxidative cyclization to thiochrome using 50 µl of a saturated solution of K<sub>3</sub>Fe(CN)<sub>6</sub> in 7 M NaOH. The oxidation reaction was neutralized after 1 min with 6 M HCl and 100 µl from the reaction mixture was analysed by reverse phase HPLC with fluorescence detection (excitation at 365 nm, emission at 450 nm). The following linear gradient, at a flow rate of 1 ml min<sup>-1</sup>, was used. Solvent A is water, solvent B is 100 mM KH<sub>2</sub>PO<sub>4</sub>, pH 6.6, and solvent C is methanol: 0 min, 100% B; 2 min, 10% A, 90% B; 10 min, 25% A, 15% B, 60% C; 12 min, 25% A, 15% B, 60% C; 15 min, 100% B;

17 min, 100% B. A calibration curve was generated using commercially available thiochrome following the same HPLC procedure.

To determine the thiamine content of the lyophilized culture medium, it was re-dissolved in 1,000 µl of double distilled H<sub>2</sub>O. This solution (500 µl) was incubated at 37 °C for 18 h with 5 mg of acid phosphatase and an aliquot (100 µl) was analysed for thiamine using the thiochrome assay described above.

The THI4/thiamine ratio was determined in quadruplicate (Fig. 4d).

**MALDI analysis of native THI4p (ScTHI4p).** The yeast cells from the large scale cultivation (6l) were harvested by centrifugation, re-suspended in the appropriate buffer and lysed by the zymolyase-homogenizer method as described above. THI4p was partially purified from the clarified protein lysate by immunomagnetic beads using Anti-His monoclonal antibody coated microbeads (µMACS, Miltenyi Biotec) following manufacturer protocol. The resulting proteins were further purified by SDS-PAGE (12%). Electrophoresis was performed at 110 V and a maximum of 15 mA for 1.5 h. The bands were stained with Coomassie blue. The protein band corresponding to THI4p (confirmed by western blot) was excised and digested with chymotrypsin at 37 °C overnight using the following protocol: the gel slice was washed with 25 mM ammonium bicarbonate (pH 8) and dehydrated with a mixture of acetonitrile (ACN) and 50 mM ammonium bicarbonate (v/v, 2:1). The washing and dehydrating steps were repeated twice. The supernatant was then removed and the gel slice was dried in a vacuum centrifuge (SpeedVac Concentrator, Savant). Chymotrypsin (Promega) dissolved in 25 mM ammonium bicarbonate (10 µl of 20 ng µl<sup>-1</sup> solution) was added to the dried gel slice. After the gel slice was completely rehydrated, 20 µl of 25 mM ammonium bicarbonate was added to cover the gel slice and proteolysis occurred at 37 °C overnight. The chymotrypsin digested sample was acidified with formic acid (pH 2–3), and desalted using a C18 ZipTip pipette tip (Millipore) according to the manufacturer's protocol. After desalting, the chymotrypsin digested sample was mixed 1:1 (v/v) with the MALDI matrix (5 mg ml<sup>-1</sup> α-cyano-4-hydroxycinnamic acid, 50% (v/v) ACN, 10 mM ammonium dihydrogen phosphate, 0.5% trifluoroacetic acid) and 1 µl of the resulting mixture was spotted onto a MALDI sample plate. MALDI-MS and MS/MS experiments were performed using a 4700 Proteomics Analyzer (Applied Biosystems). Collision-induced dissociation (CID) spectra were acquired using air as the collision gas (medium pressure setting) and at 1 kV of collision energy in the Russell laboratory at Texas A&M University.

22. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
23. DeLano, W. L. The PyMOL Molecular Graphics Systems, DeLano Scientific (2002).
24. Suter, B. *et al.* Examining protein–protein interactions using endogenously tagged yeast arrays: The cross-and-capture system. *Genome Res.* **17**, 1774–1782 (2007).
25. Ausubel, F. *et al.* *Short Protocols in Molecular Biology* 3rd edn (Wiley, 1995).

# CAREERS

**TURNING POINT** Argentinian pianist turns to biophysics PhD in the United States **p.549**

**UK RESEARCH** Industry support needed to spur innovation **p.549**

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)



The picturesque French city of Grenoble attracts both mountain climbers and ambitious researchers.

FRANCE

## Peak of potential

*Once known mostly for its natural beauty, Grenoble is becoming a centre of innovation for academia and industry.*

BY KATHARINE SANDERSON

Seated beneath some of Europe's most spectacular mountains — the Chartreuse, Vercors and Belledonne ranges in the French Alps — Grenoble is more than a playground for adrenalin-seeking skiers. It is a university and research town that hosts 22,800 research jobs, including 3,500 PhD students, 11,800 posts in public research and 7,500 in private research. Grenoble is fast becoming an innovation hub.

The city is home to one of France's longest-established science parks, and to multiple institutions intent on fostering technology transfer. The local network of angel investors is one of the most successful in the country, and Grenoble is second only to Paris in the number of French patent applications filed each year. The city has plans to expand its innovation and university links through large projects that

could bring in billions of euros in funding and create thousands of jobs. It also provides support for scientific start-ups. But there are big challenges — in particular the French academic environment, which many deem unfriendly to postdocs at best.

### INTERNATIONAL FLAIR

There are 13 public research institutes in Grenoble. Four are international: a campus of the European Molecular Biology Laboratory (EMBL); the European Synchrotron Radiation Facility; the Institute Laue–Langevin (ILL), a neutron source; and the laboratories of the Institute for Millimetre-wavelength Radioastronomy. The other institutes include the University of Grenoble, composed of multiple institutions, and French national research centres such as the Atomic and Alternative Energies Commission and the headquarters of the National Centre for Scientific Research, the

country's main research body.

The construction continues. In July, the French government announced that the Grenoble Innovation for Advanced New Technologies (GIANT) campus is to host a virtual nanoelectronics research centre that will coordinate research projects at local institutes and companies. It is expected to create 8,000 jobs, although the proportion that will be in research is not yet known. Some €460 million (US\$630 million) will be invested over 10 years, half from government and half from private investors.

"There is really a tradition of applied research and transfer to the real world here," says Véronique Pequignat, head of international business investment at the Grenoble region's economic-development body, the Agency for Studies and Promotion of Isère (AEPI). In 2009, Grenoble accounted for 780 of France's 15,000 patent applications. The city's contribution made up slightly more than one-third of the entire number of patents applied for in the surrounding Rhône-Alpes region, which includes Lyons, the second-largest metropolitan area in France.

The Inovallée, one of France's first science parks, opened in 1972 in Meylan, just outside Grenoble. It focuses on collaborations between academia and industry, and hosts 320 companies. Grenoble also provides a supportive environment for ambitious and entrepreneurial young scientists. In the past 10 years, the area has seen the creation of 200 start-ups in fields ranging from the life sciences to fluid mechanics. Of those, 132 were supported by Grenoble Alps Incubation (GRAIN), which was founded in 1999 by a group of academic and government bodies. So far, it has created more than 800 salaried research and other posts by supporting start-ups.

### INNOVATION STATION

Among the start-ups funded by GRAIN is Écrins Therapeutics, which was set up in July 2010 by Andrei Popov, a physician-turned-cancer biologist. His company has six employees — four scientists and two technicians — and is developing one anticancer product. It hopes to broaden its portfolio, says Popov, who is trying to raise €3 million to expand the business. So far, his team has received two French national innovation grants, worth €220,000 each. His company is among several located at BIOPOLIS, a biotechnology hub run by Joseph Fourier University (UJF), one of the University of Grenoble's science and technology institutions.

After being selected for investment by ►

► GRAIN, Popov was given the opportunity to attend business seminars weekly for 18 months while his idea developed. GRAIN offers its start-ups a credit line (which Popov and his team did not fully use) of up to €45,000 to be paid off through investments or money earned from sales over three years, starting three years after the company has launched. The business incubator also provides a stipend of €1,150 per month for company creators who don't have a paid job. Popov still has access to a GRAIN adviser, who offers counselling on matters from team building to financing.

Popov's experience is part of a larger trend. Science and technology are "in the ecosystem here", says Cheikhou Dieye, managing director of Grenoble Angels, a network of private investors who support small companies in their initial stages. It is among the most successful of the country's 85 such groups. When Grenoble Angels first launched in 2005, the city's tradition of established technology companies led the group to focus on semiconductor, nanotechnology and Internet companies, but biotechnology is now a big part of the portfolio; last year, four of the companies in which it invested were biotech start-ups.

### ACADEMIC AMBITIONS

Academia is also benefiting from Grenoble's science expansion. The Grenoble-Alps University of Innovation (GUI+), a super-university that will unite many of Grenoble's existing research institutions, is part of the government's Excellence Initiative (Idex) to augment university campuses. Funded by a €400-million grant, it will open as early as 2016. To hire more faculty members and staff, Yannick Vallée, head of the project and a UJF chemist, is seeking an extra €1 billion in Idex stimulus grants; he will find out in January 2012 whether he will get the funds. He estimates that the stimulus grant would pay about €40 million a year, which would fund 20 research posts and an as-yet-undetermined number of grants to support PhD students. Vallée would like to recruit researchers specializing in technology, especially nanotechnology, electronics and computer science.

Meanwhile, GIANT was selected for government funding in 2008 as one of 12 international campuses of excellence. It is being developed on a 250-hectare site that houses — and will house — some of the city's large research projects. These include MINATEC,



**"There is really a tradition of applied research and transfer to the real world here."**

Véronique Pequignat



Grenoble hosts international institutes such as the European Synchrotron Radiation Facility (centre).

a micro- and nanotechnology innovation campus that employs 2,400 researchers and files 300 patents a year. GIANT as a whole employs 6,000 researchers at the moment, and aims to grow to 10,000 researchers, 10,000 students and 10,000 industrial jobs by 2015.

However, Grenoble scientists share some frustrations with other French researchers. "Finding a job is not easy. France creates a lot of postdoc researchers and there are relatively few permanent positions," says Popov. François Briatte, a health-policy PhD student and a member of the Young Researchers' Federation (CJC) in Paris, says that academia is "in a dire state of contractual anarchy that constrains and obscures the job prospects of young researchers in France". Briatte and the CJC assert that there are "serious, enduring shortages on the tenure-track academic market". The nation's academic institutions, says Briatte, don't have the budget from central government to employ enough researchers.

### ROTATING POSITIONS

But the EMBL and the ILL offer a different employment model. The EMBL follows the same system as its locations across Europe: the Grenoble campus employs around 90 scientists, each on initial five-year contracts that are renewed for two years at a time to a maximum of nine years. The best 10% are offered permanent jobs when their contracts expire, says EMBL spokeswoman Sonia Furtado.

The ILL also has about 90 scientists on five-year contracts, but they are not offered the chance to renew. Instead, every 18 months, young scientists who are two or three years into their contracts compete for a small number of permanent positions at the centre. Over the past five years, between two and eight new scientists have been recruited each year. The arrangement, which is unique to the ILL's Grenoble location, required special permission

from the French labour council, says Elizabeth Moulin, head of recruitment and integration at the institute. In France, institutions are usually allowed to give postdocs only three-year contracts or permanent positions, she says. "The new people bring fresh blood and new ideas," says Moulin. "In general, we tell our candidates that their chance to get a permanent position at the ILL is about one in three."

Employment caveats might not be the only damper on Grenoble's scientific ambitions — a national and local distaste for animal research could also have an impact. BIOPOLIS was built on the condition that it would not contain an animal house. "For oncology research we are talking about rats and mice, the workhorses of drug discovery," says Popov. "Yet we have no right even to bring a single living mouse inside — we have to use animal houses of the academic institutes nearby. That means that the technicians have to spend an hour a day walking between BIOPOLIS and the research." Concerns about the purported risks of nanotechnology have also prompted protests: in 2005, demonstrators occupied cranes involved in building MINATEC. Briatte says that industry in Grenoble should take these protests seriously and develop a social-mediation process.

Nevertheless, the area's entrepreneurial activity remains a big draw — as do its mountains. And rather than being a distraction, some argue that the natural beauty of the area encourages a culture of ambitious scientists. "People who want to scale mountains are often successful in other areas," says Imre Berger, a genome biologist at the EMBL. "These are the people you want in your lab. They play hard and they work hard," he says. "They are eager and able to scale a pinnacle." ■

**Katharine Sanderson** is a freelance writer based in Toulouse, France.

# TURNING POINT

## Josefina del Mármol

*Argentinian PhD student Josefina del Mármol, who studies biological sciences at the Rockefeller University in New York, once planned to become a classical pianist. But Bach took a back seat, and now del Mármol is focusing on biophysics and molecular neurobiology as one of 48 inaugural Howard Hughes Medical Institute (HHMI) international student research fellows.*

### Do you miss music's former role in your life?

There's a piano at Rockefeller that everyone can use. It's so great that it's there, because I couldn't bring mine from home. I like to play all classical music. But now it's just a hobby.

### What prompted you to pursue science?

I wanted to be a musician, but in high school I took a biology class, and that changed everything. We studied evolution, and I couldn't stop reading and studying it. At college, I took up biology and never looked back.

### Describe your first major project.

As an undergraduate, I developed a fluorescent probe for tissue that lets you control what is being lit and when. That work was published last year in *Analytical Chemistry* and I was first author. It reaffirmed my interest in pursuing a science career.

### What specific area are you working in?

Mechanosensation — the conversion of mechanical stimuli into cellular responses. It has a role in the sense of touch and in physiological processes such as blood-flow sensing by vessels. Unlike for senses such as smell or vision, the molecular nature of mechanosensation remains poorly understood.

### How did you become interested in the topic?

I attended a lecture on mechanically gated ion channels by Roderick MacKinnon, and decided to do a rotation in his lab. I began monitoring how ion-channel activity responds to mechanical stimulation.

### What advice can you offer others looking to work for big names such as MacKinnon?

You have to feel genuine interest and motivation for the question being researched. If you are in it only for the prestige, it will show during your interviews.

### What has been your most significant challenge so far?

Coming to the United States. It's been a huge adjustment. All my undergraduate biology



courses were in Spanish, but here, the science is very intense, yet I have to speak and write in English. Plus it's very cold and the light gets dim at 4 o'clock in the afternoon. This will be my second winter here, and I know I'm in for months and months of suffering.

### Why did you want to study in the United States?

While I was an undergraduate in Argentina, I wasn't sure I wanted to go abroad, but I met lecturers from Rockefeller. They were so free — the way they thought, what they were doing. Scientists in Argentina are limited by money and resources, and the scientific community is much smaller than in the United States. You can't always study exactly what you want because there may not be a lab working on it. At Rockefeller, I work with top-notch scientists every day — I'm far closer to where science is actually happening.

### Has the HHMI award changed your opinion on the feasibility of a career in academia?

Yes. Coming into a graduate career in the United States, I was aware that funding is quite hard to find for international people. But so far, both Rockefeller and HHMI have supported me, which gives me the idea that it's not impossible to build up a career here regardless of my citizenship, even in times of financial crisis.

### Is there one issue that consistently crops up in your work?

Managing stress. I watch my principal investigator; although he's under a lot of pressure, he enjoys the science he does. And that's a good way to do it — be really motivated, but stay calm and try to have fun. ■

INTERVIEW BY KAREN KAPLAN

## UK RESEARCH

### Industry support falls

UK industry must invest in research and create more scientific jobs to uphold the nation's economic position, says an analysis. *Global Research Report: United Kingdom*, released on 19 October by Thomson Reuters, found that Britain is a world leader in key research indicators including highly cited papers. Almost 20% of articles with more than 1,000 citations come from the United Kingdom (that is, have at least one author in Britain), more than from any nation except the United States. Yet private investment in research has fallen since 1991. "Industry has failed to establish opportunities for talented researchers," says report author Jonathan Adams, Thomson Reuters' director of research evaluation, based in Leeds, UK. "We're going to find ourselves heading for second-rate economic status."

## AUSTRALIA

### Academics unhappy

Australian academic researchers are rallying behind a report that laments their working conditions. The government-funded study, out in September, surveyed 5,525 academics across all career stages and fields at 20 universities. It found that nearly half of academics under 30 want to leave the country or the profession owing to low pay and lack of job security. Researchers are frustrated by teaching obligations that cut into research time; low grant success rates; and 70- to 80-hour working weeks. Emmaline Bexley, a lecturer in higher education at the University of Melbourne and lead author of the study, says she hopes that her research will "help government and universities to work together to replenish the academic workforce".

## COLLABORATION

### Regional pact formed

An agreement will let postdocs and early-career researchers from Singapore and Europe apply for training funds in each other's regions. Under the three-year pact, which was announced on 13 October and aims to stimulate collaboration, Singaporean scientists can seek European Molecular Biology Organization (EMBO) fellowships and grants and undertake EMBO training courses and activities. Fellowships will be available to European scientists wishing to work in Singapore. The pact is between EMBO, the European Molecular Biology Conference and the Singaporean government.

# AN EASY SALE

*Time for an upgrade.*

BY BIREN SHAH

When the perky birthsuiter girl came to greet him, Allen's EPU — emotion processing chip — interrupted normal conscious processing and adjusted internal state variables to disappointment.

"Hi, I'm Jeannette," the girl said with a big smile.

"Are you sure we can't do this over the net?" Allen asked.

"Don't be silly. You're already here." She turned on one high heel.

Allen centred his visual focus below the edge of her black pinstriped skirt-suit, watching the smooth skin of her calves ripple as she walked into the conference room. He instructed *follow*, and a subroutine activated his leg actuators. After a few nanoseconds, his positional readings began to change. His visual cortex reallocated resources to handle the optic flow, shrinking the area he could see in focus. *It's just like them, he thought. They don't know a damned thing about their customers.*

When he finally made it into the chair across from her, she said: "You've been a vocal critic of most of our releases since the 116 series you're wearing now —"

He began queuing up his response. "You keep introducing upgrades for huge profits and keep gouging your customers..."

As his voice modulator processed and spoke his words, Jeannette flipped her lush black hair and rubbed the brown skin of her neck with the back of her hand. Her tendons pulled tight, framing the hollow of her neck in a little v.

If they were jacked in, he'd be the one waiting on her. And there certainly wouldn't be any distractions like *this*.

His EPU launched the memory of his last touch. He and his wife had been outside, sitting on a lawn. The sun warmed his shrivelled flesh. A weak breeze cooled his left arm from behind. She sat between his legs, leaning against him as his atrophied muscles strained to

hold their weight. He'd insisted he get to be the one to hold her despite her protests that he was too weak from the chemo. It had been his last chance to feel like a true man. He leaned closer, just by having the desire, and her blond hair tickled his ear as it fluttered against him. He took a deep breath to commit the smell of her — vanilla and jasmine and a touch of sweat — to memory. His fingers stuck to her glistening pale skin as he caressed her arm, then slipped only to catch again. Goosebumps rose where he had touched.

It was one of his last birthsuit memories, stored perfectly in the artificial brain housing his mind. He could go back to that moment and re-experience every sensation. He never did.

"I'm familiar with your objections," Jeannette said.

Allen realized he'd just trailed off. This beautiful girl had him swimming in those terrible currents of sensation again. He shut down the memory-playing process — but he had to wait for his emotion chip to flush his state buffer of the downward accelerometer reading meant to simulate a sinking feeling of loss. "Then you know I've added every feature you released since the 116s aftermarket for a tenth of the cost of an upgrade."

"I'm not here to argue that with you." She leaned forward and Allen couldn't pull his visual focus from a flash of cobalt blue satin where her jacket parted.

"Enjoying the view?" she asked.

The salmon-coloured chromoplasts in Allen's cheeks flared. His central clock rate accelerated.

"Don't worry — it's kind of why we're here." Jeannette stood up and removed her jacket. "Although you're right that most customers got only aesthetic value out of the 200s, 300s and 400s, those models laid the groundwork for a huge leap forward in our

product map." She turned around. The blue blouse cut down to her waist, exposing the unblemished skin of her back. She pressed against her neck and a small flap appeared. She pulled on it gently, peeling the skin away to reveal a titanium alloy spine.

"That's not a birthsuit," Allen struggled to get the words queued up between his EPU's rapid-fire interrupts.

"No, you're getting an advanced look at the 500 — all the technologies we've been developing for the past 150 years wrapped in a biological prosthetic." She pushed her skin back in place and sat down.

Allen focused on her forearm and instructed *index-finger forward-back, forward-back*. His finger reached out and poked her twice. "Does this feel... right?"

"The 500 is virtually indistinguishable from a birthsuit, from inside and out."

Allen focused on her hand and instructed *palm surface dock (tolerance: 30 cm of current position) followed by 5 cm forward-back, 15 seconds*. His hand rubbed the back of her hand. The sensory feedback came back like he was rubbing a pan with a spatula.

"We want to offer you an upgrade, free of charge."

Manipulation-warning subroutines fired, but Allen dismissed them and shut down the monitoring process. *What did she feel like? Smell like?*

"Your blog reaches nearly 10 million of our most savvy customers. As great as the 500s are, after an initial exploring phase, our beta testers regularly end up weighing over 300 pounds. Obviously, this causes some dissatisfaction."

Allen zoomed his visual field in on the shape-shifting of her mouth as she talked. He tried to remember what it was to have such a mouth — supple, responsive...

"We want you in our PR department, educating our customers on the maintenance requirements of a biological prosthesis. We're thinking for say a period of 50 years?"

A contract appeared in Allen's inbox. With a single thought, he signed the document with his public key and sent it back. "What was that about dissatisfaction?"

"Never mind," she answered after a screen in the table flashed a message. "You won't have the same problems. I can tell." ■

**Biren Shah** has recently moved to New York City, changed careers, and started writing again. He's exploring as fast as he can. You can watch at [www.birenshah.com](http://www.birenshah.com).



► **NATURE.COM**  
Follow Futures on  
Facebook at:  
[go.nature.com/mtoodm](http://go.nature.com/mtoodm)

# Verbal and non-verbal intelligence changes in the teenage brain

Sue Ramsden<sup>1</sup>, Fiona M. Richardson<sup>1</sup>, Goulven Josse<sup>1</sup>, Michael S. C. Thomas<sup>2</sup>, Caroline Ellis<sup>1</sup>, Clare Shakeshaft<sup>1</sup>, Mohamed L. Seghier<sup>1</sup> & Cathy J. Price<sup>1</sup>

Intelligence quotient (IQ) is a standardized measure of human intellectual capacity that takes into account a wide range of cognitive skills<sup>1</sup>. IQ is generally considered to be stable across the lifespan, with scores at one time point used to predict educational achievement and employment prospects in later years<sup>1</sup>. Neuroimaging allows us to test whether unexpected longitudinal fluctuations in measured IQ are related to brain development. Here we show that verbal and non-verbal IQ can rise or fall in the teenage years, with these changes in performance validated by their close correlation with changes in local brain structure. A combination of structural and functional imaging showed that verbal IQ changed with grey matter in a region that was activated by speech, whereas non-verbal IQ changed with grey matter in a region that was activated by finger movements. By using longitudinal assessments of the same individuals, we obviated the many sources of variation in brain structure that confound cross-sectional studies. This allowed us to dissociate neural markers for the two types of IQ and to show that general verbal and non-verbal abilities are closely linked to the sensorimotor skills involved in learning. More generally, our results emphasize the possibility that an individual's intellectual capacity relative to their peers can decrease or increase in the teenage years. This would be encouraging to those whose intellectual potential may improve, and would be a warning that early achievers may not maintain their potential.

An individual's abilities and capacity to learn can be partly captured by the use of verbal and non-verbal (henceforth performance) intelligence tests. IQ provides a standardized method for measuring intellectual abilities and is widely used within education, employment and clinical practice. In the absence of neurological insult or degenerative conditions, IQ is usually expected to be stable across lifespan, as evidenced by the fact that IQ measurements made at different points in an individual's life tend to correlate well<sup>1,2</sup>. Nevertheless, strong correlations over time disguise considerable individual variation; for example, a correlation coefficient of 0.7 (which is not unusual with verbal IQ) still leaves over 50% of the variation unexplained. The study that we report here tested whether variation in a teenager's IQ over time correlated with changes in brain structure. If it did, this would provide construct validity for the increase or decrease of IQ in the teenage years, because if IQ changes correspond to structural brain changes then they are unlikely to represent measurement error in the IQ tests. In addition, if verbal and performance skills change at different rates in different individuals, the neural markers for verbal and performance IQ changes could in principle be dissociated. This would overcome two of the challenges faced by previous studies of between-subject variability in IQ measures at a given time point: verbal and performance IQ are tightly correlated in individuals, so it has been hard to identify neural structures corresponding to each<sup>3,4</sup>; and there are many sources of between-subject variance in brain structure (for example gender, age, size and handedness) that hide the relevant differences.

Our participants were 33 healthy and neurologically normal adolescents with a deliberately wide and heterogeneous mix of abilities (see Supplementary Information for details and the implications of our sampling for the generalizability of our conclusions). They were first tested in 2004 ('time 1') when they were 12–16 yr old (mean, 14.1 yr). Testing was repeated in 2007/2008 ('time 2') when the same individuals were 15–20 yr old (mean, 17.7 yr). See Table 1 for further details of the participants. During the intervening years, there were no testing sessions, and participants and their parents had no knowledge that they would be invited back for further testing. On both test occasions, each participant had a structural brain scan using magnetic resonance imaging (MRI) and had their IQ measured using the Wechsler Intelligence Scale for Children (WISC-III) at time 1 and the Wechsler Adult Intelligence Scale (WAIS-III) at time 2 (see Supplementary Information for details). These two widely used, age-appropriate assessments<sup>5</sup> produce strongly correlated results at a given time point, consistent with them measuring highly similar constructs<sup>6</sup>. Scores on individual subtests are standardized against age-specific norms and then grouped to produce separate measures of verbal IQ (VIQ) and performance IQ (PIQ), with VIQ encompassing those tests most related to verbal skills and PIQ being more independent of verbal skills. Nevertheless, VIQ and PIQ scores are very significantly correlated with each other across participants: in our sample, the correlations between VIQ and PIQ were  $r = 0.51$  at time 1 and  $r = 0.55$  at time 2 (in both cases,  $n = 33$ ;  $P < 0.01$ ). Full-scale IQ (FSIQ) is the composite of VIQ and PIQ and is regarded as the best measure of general intellectual capacity (the  $g$  factor) that has previously been shown to correlate with brain size and cortical thickness in a wide variety of frontal, parietal and temporal brain regions<sup>7,8</sup>.

The wide range of abilities in our sample was confirmed as follows: FSIQ ranged from 77 to 135 at time 1 and from 87 to 143 at time 2, with averages of 112 and 113 at times 1 and 2, respectively, and a tight correlation across testing points ( $r = 0.79$ ;  $P < 0.001$ ). Our interest was in the considerable variation observed between testing points at the individual level, which ranged from  $-20$  to  $+23$  for VIQ,  $-18$  to  $+17$  for PIQ and  $-18$  to  $+21$  for FSIQ. Even if the extreme values of the published 90% confidence intervals are used on both occasions, 39% of the sample showed a clear change in VIQ, 21% in PIQ and 33% in FSIQ. In terms of the overall distribution, 21% of our sample showed

**Table 1 | Participants' details**

	Datum	Age	FSIQ	VIQ	PIQ
Time 1	Mean (s.d.)	14.1 (1.0)	112 (13.9)	113 (15.1)	108 (12.3)
	Min, max	12.6, 16.5	77, 135	84, 139	74, 137
Time 2	Mean (s.d.)	17.7 (1.0)	113 (14.0)	116 (18.0)	107 (9.6)
	Min, max	15.9, 20.2	87, 143	90, 150	83, 124
Correlation*	$r$	—	0.792†	0.809†	0.589†
Change	Mean (s.d.)	3.5 (0.2)	+1.0 (9.0)	+3.0 (10.6)	-1.0 (10.2)
	Min, max	3.3, 3.9	-18, +21	-20, +23	-18, +17

\* Correlation coefficient between scores at times 1 and 2. †  $P < 0.01$ .

$n = 33$  (19 male, 14 female). s.d., standard deviation.

<sup>1</sup>Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, UK. <sup>2</sup>Developmental Neurocognition Laboratory, Department of Psychological Sciences, Birkbeck College, University of London, London WC1E 7HX, UK.

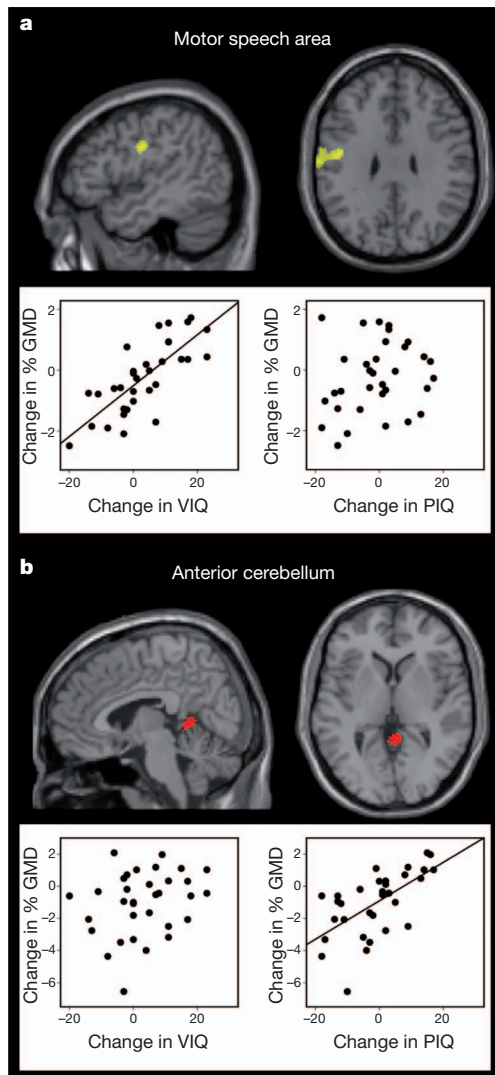
a shift of at least one population standard deviation (15) in the VIQ measure, and 18% in the PIQ measure. However, only one participant had a shift of this magnitude in both measures, and, for that participant, one measure showed an increase and the other a decrease. This pattern is reflected in the absence of a significant correlation between the change in VIQ and the change in PIQ. The independence of

changes in these two measures allows us to investigate the effect of each without confounding influences from the other.

To test whether the observed IQ changes were meaningfully reflected in brain structure, we correlated them with changes in local brain structure. This within-subject correlation obviates the many possible sources of between-subject variance and may have sensitized our analysis to neural markers of VIQ and PIQ that have not previously been revealed. Given the distributed nature of brain regions associated with between-subject differences in FSIQ<sup>7–9</sup>, regions of interest were not used in this analysis, and the results of the whole-brain analysis were only considered to be significant at  $P < 0.05$  after familywise error correction in either height (peak signal at a single voxel) or extent (number of voxels that were significant at  $P < 0.001$ ).

Using regression analysis, we studied the brain changes associated with a change in VIQ, PIQ or FSIQ (see Methods Summary for details). The results (Fig. 1) showed that changes in VIQ were positively correlated with changes in grey matter density (and volume) in a region of the left motor cortex that is activated by the articulation of speech<sup>10</sup>. Conversely, changes in PIQ were positively correlated with grey matter density in the anterior cerebellum (lobule IV), which is associated with motor movements of the hand<sup>11,12</sup>. *Post hoc* tests that correlated structural change with change in each of the nine VIQ and PIQ subtest scores that were common in the WISC and WAIS assessments found that the neural marker for VIQ indexed constructs that were shared by all VIQ measures and that the neural marker for PIQ indexed constructs that were common to three of the four PIQ measures (Table 2). This indicates that our VIQ and PIQ markers indexed skills that were not specific to individual subtests. There were no other grey or white matter effects that reached significance in a whole-brain structural analysis of VIQ, PIQ or FSIQ. See Supplementary Information for details of further *post hoc* tests.

Our findings that VIQ changes were related to a motor speech region and PIQ changes were related to a motor hand region are consistent with previous claims that cognitive intelligence is partly dependent on sensorimotor skills<sup>13–18</sup>. Using functional imaging in the same 33 participants performing a range of sensory, motor and language tasks, we confirmed that the left motor speech region identified in the VIQ structural analysis was more activated by articulation tasks (naming, reading and saying “one, two, three”) than by semantic or perceptual tasks that required a finger press response (see Supplementary Information for details). In contrast, a region very close to the anterior cerebellum region identified in the PIQ structural analysis was more activated during tasks involving finger presses than during tasks involving articulation. Figure 2 shows these results at both the group level and the individual level. The locations of the grey



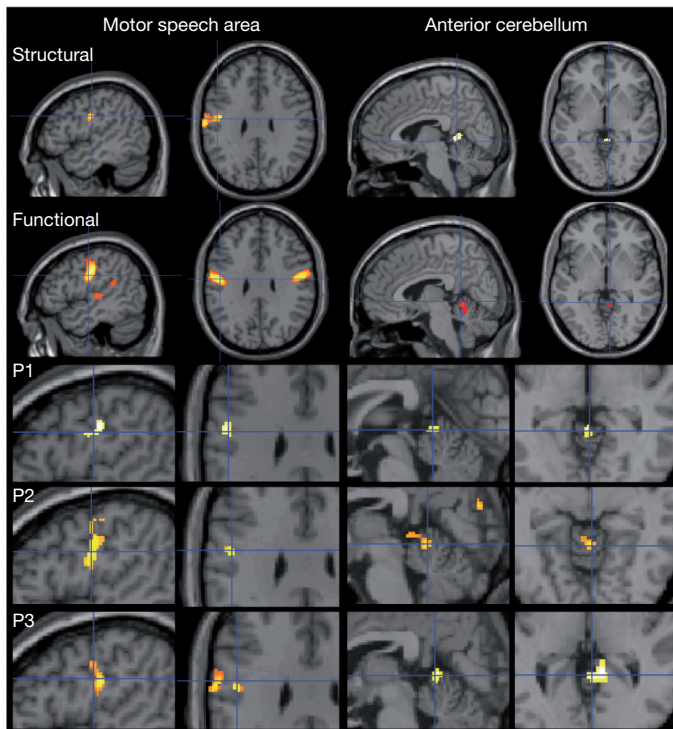
**Figure 1 | Location of brain regions where grey matter changed with VIQ and PIQ.** **a**, Correlation between change in grey matter density and change in VIQ (yellow) in the left motor speech region (peak in the left precentral gyrus at  $x = -47$  mm,  $y = -9$  mm,  $z = +30$  mm, measured in Montreal Neurological Institute (MNI) space, with a Z-score of 5.2 and 681 voxels at  $P < 0.001$ ). The corresponding effect on volume was slightly less significant (Z-score, 3.5; 118 voxels at  $P < 0.001$ ). **b**, Correlation between change in PIQ (red) and change in grey matter density in the anterior cerebellum (peak at  $x = +6$  mm,  $y = -46$  mm,  $z = +3$  mm, in MNI space, with a Z-score of 3.9 and 210 voxels at  $P < 0.001$ ). Both effects were significant at  $P < 0.05$  after familywise error correction for multiple comparisons in extent based on the number of voxels in a cluster that survived  $P < 0.001$  uncorrected. In addition, the VIQ effect was significant at  $P < 0.05$  after familywise error correction for multiple comparisons in height. The statistical threshold used in the figure ( $P < 0.001$ ) illustrates the extent of the effects. Plots show the change in grey matter density versus the change in both VIQ and PIQ at the voxel with the highest Z-score in the appropriate region. Linear regression lines are shown for significant correlations. Changes in the motor speech region correlated with changes in VIQ but not changes in PIQ, whereas changes in the anterior cerebellum correlated with changes in PIQ but not changes in VIQ ( $P < 0.001$ ).  $n = 33$ ; GMD, grey matter density.

**Table 2 | Correlations between grey matter density and score**

Test type		Motor speech region ( $r$ )	Anterior cerebellum ( $r$ )
Verbal tests	Vocabulary	0.284*	0.142
	Similarities	0.438†	-0.021
	Arithmetic	0.477†	0.304‡
	Information	0.314‡	0.185
	Comprehension	0.541†	0.183
Non-verbal tests	Picture completion	0.038	0.363‡
	Digit symbol coding	0.003	-0.028
	Block design	0.000	0.306‡
	Picture arrangement	0.126	0.437†

\*Trend (one-tailed) at  $P = 0.0545$ . †Significant (one-tailed) at  $P < 0.01$ . ‡Significant (one-tailed) at  $P < 0.05$ .

Correlations were calculated using changes in scaled (that is, age-adjusted) scores in the various subtests that were common to both the WISC and the WAIS. The change in grey matter density in the motor speech region correlated significantly with changes in scores in four of the five verbal subtests, and there was a near-significant trend in the fifth but it did not correlate significantly with changes in scores in any of the four tests that comprise PIQ. Conversely, the change in grey matter density in the anterior cerebellum correlated significantly with changes in scores in three of the four tests that comprise PIQ (the exception being the digit symbol coding test, which has a particular loading on processing speed) but correlated with changes in scores in only one of the verbal tests (the arithmetic test, which probably has the smallest verbal component of the verbal tasks).



**Figure 2 | Functional activations in the regions identified by the structural analysis.** The motor speech region was more activated by articulation tasks than by finger press tasks ( $x = -48$  mm,  $y = -10$  mm,  $z = +30$  mm (MNI);  $t = 14.7$ ;  $P < 0.05$  familywise-error-corrected for multiple comparisons across the whole brain), and corresponds to the region identified in the structural analysis for VIQ. These effects were consistently observed at the same coordinates for all individual subjects. In the three exemplar participants shown here (P1, P2, P3), the Z-scores were 3.9, 3.5 and 3.0, respectively. The anterior cerebellum region was more activated during finger presses than during articulation at both the group level (peak at  $x = +6$  mm,  $y = -48$  mm,  $z = -4$  mm (MNI); Z-score, 3.7; 216 voxels at  $P < 0.001$  corrected for multiple comparisons in extent) and the individual level (P1:  $x = +12$  mm,  $y = -48$  mm,  $z = +2$  mm (MNI); Z-score, 3.7; P2:  $x = +6$  mm,  $y = -50$  mm,  $z = -6$  mm (MNI); Z-score, 3.3; P3:  $x = +12$  mm,  $y = -46$  mm,  $z = +2$  mm (MNI); Z-score, 4.9). In all cases, the activation peaks were identified from whole-brain analyses and the peak effects for the correlation with structure are illustrated with blue cross hairs in both the structural results and the functional results. This illustrates that the location of the structural effects is within the regions identified by the functional effects.

matter changes associated with VIQ and PIQ changes do not correspond to the anterior frontal and parietal regions associated with general intelligence<sup>7</sup> ( $g$  factor). It may therefore be the case that  $g$  remains relatively constant across ages, but changes in the ability to perform individual subtests depend on changes in sensorimotor skills. It is also notable that although completion of the subtests comprising verbal and performance measures must implicate a network of brain regions, only structural changes in regions associated with sensorimotor skills showed correlations with changes in VIQ and PIQ.

The changes in brain structure that correlated with changes in IQ allow us to explain some of the variance in terms of brain development. Specifically, 66% of the variance in VIQ at time 2 was accounted for by VIQ at time 1, a further 20% was accounted for by the change in grey matter density in the left motor speech region, with the remaining 14% unaccounted for. Similarly, 35% of the variance in PIQ at time 2 was accounted for by PIQ at time 1, with 13% accounted for by the change in grey matter density in the anterior cerebellum, leaving 52% unaccounted for. Future studies may be able to account for more of the between-subject variability by using a similar methodology with larger samples or other methodologies that measure structural or functional connectivity<sup>8,19</sup>.

Our findings demonstrate considerable effects of brain plasticity in our sample during the teenage years, over and above normal development. By obviating the many sources of between-subject variance and controlling for global changes in brain structure, our within-subject analysis has allowed us to dissociate brain regions where structure reflects individual differences in verbal or non-verbal performance, in a way that has proved difficult in previous studies using behavioural data from a single point in time. We have also shown that the changes observed over time in the IQ scores of teenagers cannot simply be measurement error, because they correlate with independently measured changes in brain structure in regions that are plausibly related to the verbal and non-verbal functions tested. Further studies are required to determine the generalizability of this finding; for example, the same degree of plasticity may be present throughout life or the adolescent years covered by this study may be special in this regard. In addition, future work could consider the causes of the identified changes both in intelligence and in brain structure and how they impact on educational performance and employment prospects. The implication of our present findings is that an individual's strengths and weaknesses in skills relevant to education and employment are still emerging or changing in the teenage years.

## METHODS SUMMARY

This study was approved by the Joint Ethics Committee of the Institute of Neurology and the National Hospital for Neurology and Neurosurgery, London, UK. All structural and functional scans at times 1 and 2 were acquired from the same Siemens 1.5T Sonata MRI scanner (Siemens Medical Systems). The structural images were acquired using a T1-weighted modified driven equilibrium Fourier transform sequence with 176 sagittal partitions and an image matrix of  $256 \times 224$ , yielding a final resolution of  $1 \text{ mm}^3$  (repetition time, 12.24 ms; echo time, 3.56 ms; inversion time, 530 ms). To pre-process the 66 structural images (33 participants  $\times$  2 time points), we used SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>) with the DARTEL toolbox to segment and spatially normalize the brains into the same template, with and without modulation. Modulated images incorporate a measure of local brain volume, whereas unmodulated images, used with proportional scaling to correct for global grey matter, provide a measure of regional grey matter density. Previous studies<sup>20–22</sup> have shown that the correlations between brain structure and cognitive ability are better detected by grey matter density. Coordinates for each voxel were converted to standard MNI space. Images were smoothed using a Gaussian kernel with an isotropic full-width of 8 mm at half-maximum. The relationship between change in IQ and change in brain structure was investigated by entering the appropriate pre-processed images (modulated or unmodulated grey or white matter) into within-subject paired  $t$ -tests, with change in IQ (VIQ, PIQ or FSIQ) and year of scan as covariates. The degree to which IQ at time 2 was predicted by changes in brain structure was investigated in a hierarchical regression analysis with IQ at time 1 entered before change in brain structure. Details of the functional imaging method have been reported elsewhere<sup>23–25</sup> and are summarized in Supplementary Information.

Received 17 May; accepted 26 August 2011.

Published online 19 October 2011.

- McCall, R. B. Childhood IQs as predictors of adult educational and occupational status. *Science* **197**, 482–483 (1977).
- Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R. & Starr, J. M. The stability of differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. *Intelligence* **28**, 49–55 (2000).
- Wilke, M., Sohn, J.-H., Byars, A. W. & Holland, S. K. Bright spots: correlations of gray matter volume with IQ in a normal pediatric population. *Neuroimage* **20**, 202–215 (2003).
- Gong, Q.-Y. *et al.* Voxel-based morphometry and stereology provide convergent evidence of the importance of medial prefrontal cortex for fluid intelligence in healthy adults. *Neuroimage* **25**, 1175–1186 (2005).
- Camara, W. J., Nathan, J. S. & Puente, A. E. Psychological test usage: implications in professional psychology. *Prof. Psychol. Res. Pr.* **31**, 141–154 (2000).
- Kaufman, A. & Lichtenberger, E. O. *Assessing Adolescent and Adult Intelligence* 209–216 (Wiley, 2006).
- Haier, R. J., Jung, R. E., Yeo, R. A., Head, K. & Alkire, M. T. Structural brain variation and general intelligence. *Neuroimage* **23**, 425–433 (2004).
- Colom, R., Karama, S., Jung, R. E. & Haier, R. J. Human intelligence and brain networks. *Dialogues Clin. Neurosci.* **12**, 489–501 (2010).
- Shaw, P. *et al.* Intellectual ability and cortical development in children and adolescents. *Nature* **440**, 676–679 (2006).
- Huang, J., Carr, T. H. & Cao, Y. Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum. Brain Mapp.* **15**, 39–53 (2002).

11. Nitschke, M. F., Kleinschmidt, A., Wessel, K. & Frahm, J. Somatotopic motor representation in the human anterior cerebellum: a high-resolution functional MRI study. *Brain* **119**, 1023–1029 (1996).
12. Stoodley, C. J., Valerad, E. M. & Schmahmann, J. D. An fMRI study of intra-individual functional topography in the human cerebellum. *Behav. Neurol.* **23**, 65–79 (2010).
13. Diamond, A. Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Dev.* **71**, 44–56 (2000).
14. Pangelinan, M. M. *et al.* Beyond age and gender: relationships between cortical and subcortical brain volume and cognitive-motor abilities in school-age children. *Neuroimage* **54**, 3093–3100 (2011).
15. Davis, A. S., Pass, L. A., Finch, W. H., Dean, R. S. & Woodcock, R. W. The canonical relationship between sensory-motor functioning and cognitive processing in children with attention-deficit/hyperactivity disorder. *Arch. Clin. Neuropsychol.* **24**, 273–286 (2009).
16. Davis, E. E., Pitchford, N. J., Jaspan, T., McArthur, D. & Walker, D. Development of cognitive and motor function following cerebellar tumour injury sustained in early childhood. *Cortex* **46**, 919–932 (2010).
17. Rosenbaum, D. A., Carlson, R. A. & Gilmore, R. O. Acquisition of intellectual and perceptual-motor skills. *Annu. Rev. Psychol.* **52**, 453–470 (2001).
18. Wassenberg, R. *et al.* Relation between cognitive and motor performance in 5- to 6-year-old children: results from a large-scale cross-sectional study. *Child Dev.* **76**, 1092–1103 (2005).
19. Jung, R. E. & Haier, R. J. The parieto-frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav. Brain Sci.* **30**, 135–154 (2007).
20. Eckert, M. A. *et al.* To modulate or not to modulate: differing results in uniquely shaped Williams syndrome brains. *Neuroimage* **32**, 1001–1007 (2006).
21. Lee, H. *et al.* Anatomical traces of vocabulary acquisition in the adolescent brain. *J. Neurosci.* **27**, 1184–1189 (2007).
22. Richardson, F. M., Thomas, M. S., Filippi, R., Harth, H. & Price, C. J. Contrasting effects of vocabulary knowledge on temporal and parietal brain structure across lifespan. *J. Cogn. Neurosci.* **22**, 943–954 (2010).
23. Seghier, M. L., Fagan, E. & Price, C. J. Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *J. Neurosci.* **30**, 16809–16817 (2010).
24. Seghier, M. L. & Price, C. J. Dissociating functional brain networks by decoding the between-subject variability. *Neuroimage* **45**, 349–359 (2009).
25. Parker Jones, 'Ö. *et al.* Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb. Cortex* advance online publication, (<http://dx.doi.org/10.1093/cercor/bhr161>) (24 June 2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was funded by the Wellcome Trust. We thank J. Glensman, A. Brennan, A. Peters, L. Stewart, K. Pitcher and R. Rutherford for their help with data collection; and W. Penny for his advice on statistical analyses.

**Author Contributions** C.J.P. designed and supervised the study. C.J.P. and C.S. recruited the participants. C.S., S.R. and G.J. collected the data. F.M.R., S.R., C.E., M.L.S. and C.J.P. analysed the data. S.R., M.S.C.T. and C.J.P. wrote the manuscript and all authors edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to C.J.P. ([c.price@fil.ion.ucl.ac.uk](mailto:c.price@fil.ion.ucl.ac.uk)).

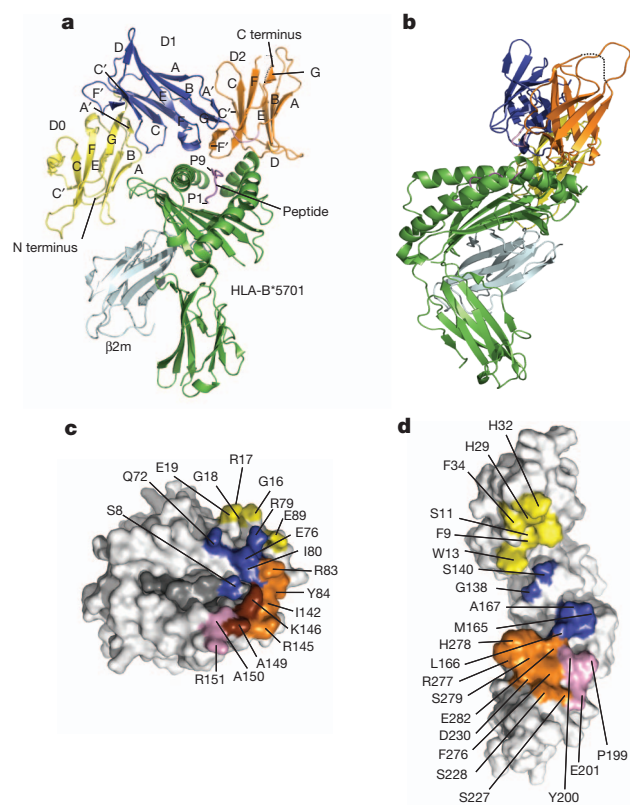
# Killer cell immunoglobulin-like receptor 3DL1-mediated recognition of human leukocyte antigen B

Julian P. Vivian<sup>1</sup>, Renee C. Duncan<sup>1</sup>, Richard Berry<sup>1</sup>, Geraldine M. O'Connor<sup>2</sup>, Hugh H. Reid<sup>1</sup>, Travis Beddoe<sup>1</sup>, Stephanie Gras<sup>1</sup>, Philippa M. Saunders<sup>3</sup>, Maya A. Olshina<sup>1</sup>, Jacqueline M. L. Widjaja<sup>3</sup>, Christopher M. Harpur<sup>3</sup>, Jie Lin<sup>3</sup>, Sebastien M. Maloeste<sup>4</sup>, David A. Price<sup>5,6</sup>, Bernard A. P. Lafont<sup>4</sup>, Daniel W. McVicar<sup>2</sup>, Craig S. Clements<sup>1</sup>, Andrew G. Brooks<sup>3</sup> & Jamie Rossjohn<sup>1,5</sup>

Members of the killer cell immunoglobulin-like receptor (KIR) family, a large group of polymorphic receptors expressed on natural killer (NK) cells, recognize particular peptide-laden human leukocyte antigen (pHLA) class I molecules and have a pivotal role in innate immune responses<sup>1</sup>. Allelic variation and extensive polymorphism within the three-domain KIR family (KIR3D, domains D0–D1–D2) affects pHLA binding specificity and is linked to the control of viral replication and the treatment outcome of certain haematological malignancies<sup>1–3</sup>. Here we describe the structure of a human KIR3DL1 receptor bound to HLA-B\*5701 complexed with a self-peptide. KIR3DL1 clamped around the carboxy-terminal end of the HLA-B\*5701 antigen-binding cleft, resulting in two discontinuous footprints on the pHLA. First, the D0 domain, a distinguishing feature of the KIR3D family, extended towards  $\beta$ 2-microglobulin and abutted a region of the HLA molecule with limited polymorphism, thereby acting as an ‘innate HLA sensor’ domain. Second, whereas the D2–HLA-B\*5701 interface exhibited a high degree of complementarity, the D1–pHLA-B\*5701 contacts were suboptimal and accommodated a degree of sequence variation both within the peptide and the polymorphic region of the HLA molecule. Although the two-domain KIR (KIR2D) and KIR3DL1 docked similarly onto HLA-C<sup>4,5</sup> and HLA-B respectively, the corresponding D1-mediated interactions differed markedly, thereby providing insight into the specificity of KIR3DL1 for discrete HLA-A and HLA-B allotypes. Collectively, in association with extensive mutagenesis studies at the KIR3DL1–pHLA-B\*5701 interface, we provide a framework for understanding the intricate interplay between peptide variability, KIR3D and HLA polymorphism in determining the specificity requirements of this essential innate interaction that is conserved across primate species.

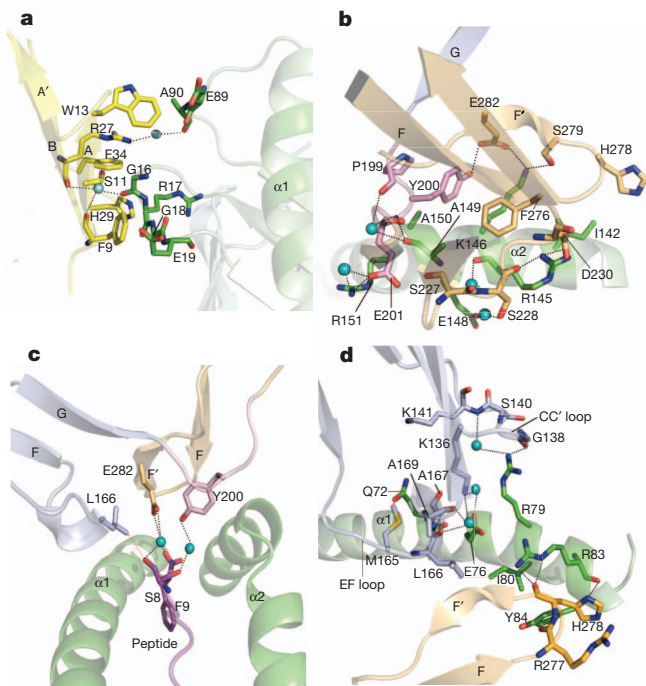
HLA-B57 carriage has been associated with delayed progression to AIDS in HIV-infected individuals, with a strong genetic association between the KIR3DL1–HLA-B57 interaction, reduced viral loads and delayed HIV disease progression<sup>3</sup>. We expressed KIR3DL1\*001, a prototypical family member, and co-complexed it with HLA-B\*5701 bound to a self-peptide (LSSPVTKSF). The affinity ( $K_D$ ) of this interaction was approximately 17  $\mu$ M (Supplementary Table 1 and Supplementary Fig. 1). We then determined the KIR3DL1\*001–HLA-B\*5701–LSSPVTKSF structure to 1.8 Å resolution (Supplementary Table 2 and Supplementary Fig. 2). KIR3DL1\*001 clamped around the C-terminal end of the HLA-B\*5701 antigen-binding cleft (Fig. 1a, b), forming an extensive interface (total buried surface area (BSA), 1,740 Å<sup>2</sup>) that encompassed two discontinuous sites—one mediated via the D0 domain and the other via the D1–D2 domains (Figs 1c, d and 2a–d). KIR3DL1\*001 adopted an elongated, zigzag conformation, with the three immunoglobulin (Ig) domains, termed D0, D1 and D2

(residues 7–98, 99–198 and 203–292, respectively) defined by the E-type Ig fold topology (Fig. 1a). The D0 domain, a feature of the KIR3D family<sup>6</sup> packed against the D1 domain, the relative juxtapositioning of which (83°) is similar to that of the D1–D2 inter-domain angle (81°), which in turn is analogous to the relative orientation of D1–D2 domains (76°) found in the KIR2D receptors (root mean squared deviation (r.m.s.d.) of D1–D2 domains in KIR2DL1 and



**Figure 1 | Structure of the KIR3DL1\*001–pHLA-B\*5701 complex.** a, b, Orthogonal views of the complex with the KIR3DL1\*001  $\beta$ -strands labelled. The HLA and  $\beta$ 2-microglobulin ( $\beta$ 2m) are coloured green and cyan, respectively; D0, D1, D1–D2 loop and D2 are coloured yellow, blue, pink and orange, respectively; dashed line represents the unresolved loop between the E and F  $\beta$ -strands. c, d, The footprint mapped to the surface of HLA and KIR3DL1\*001, respectively, with residues coloured in each case according to the interacting KIR3DL1\*001 domain: D0 (yellow), D1 (blue), D1–D2 loop (pink) and D2 (orange). Residues that contact the linker and the D2 domain are coloured brown.

<sup>1</sup>Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia. <sup>2</sup>Cancer and Inflammation Program, National Cancer Institute-Frederick, Frederick, Maryland 21702, USA. <sup>3</sup>Department of Microbiology & Immunology, University of Melbourne, Parkville, Victoria 3010, Australia. <sup>4</sup>Non-Human Primate Immunogenetics and Cellular Immunology Unit, Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>5</sup>Department of Infection, Immunity and Biochemistry, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, Wales, UK. <sup>6</sup>Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.



**Figure 2 | Contacts between the KIR3DL1\*001 receptor and pHLA-B\*5701.** Panels coloured as in Fig. 1. Waters are shown as cyan spheres; hydrogen bonds as black lines; van der Waals contacts as red lines. **a**, Contacts between the D0 domain and pHLA. **b**, Contacts between the receptor and the  $\alpha 2$  helix of pHLA. **c**, Contacts to the peptide. Residues from the D1 and D2 domains form a single van der Waals and three water-mediated contacts to the P8 and P9 peptide positions. **d**, Contacts between the receptor and the  $\alpha 1$  helix. The interface between the D1 domain and the  $\alpha 1$  helix was suboptimal, comprising a single direct hydrogen bond from Gly 138 to Arg 79.

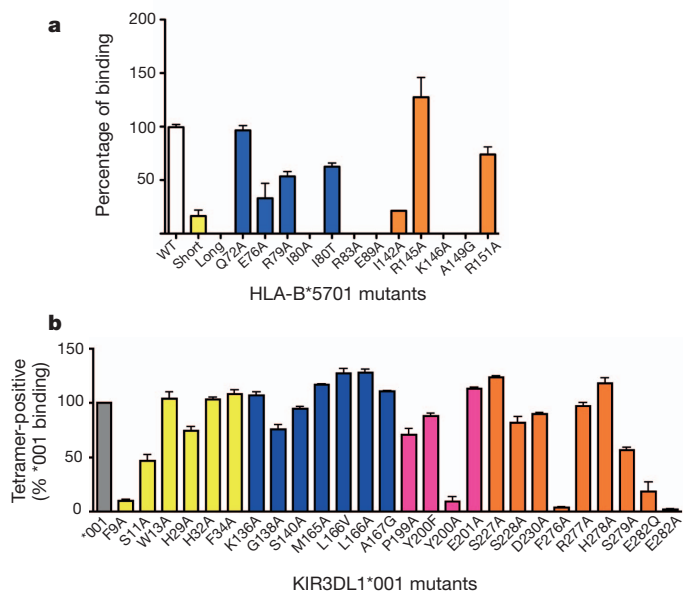
KIR3DL1 is 1.2 Å (Supplementary Fig. 3a)<sup>4,5</sup>. Further, the positioning of the D0 domain relative to the D1 and D2 domains appears to be fixed (Supplementary Fig. 3b, c), thereby generating a pre-formed pHLA-binding receptor.

The D0 domain contributed 30% BSA upon complexation with ligand, being orientated almost perpendicular to the main axis of the antigen-binding cleft, extending towards, and just contacting,  $\beta 2$ -microglobulin (Fig. 1a). A surface-exposed aromatic cluster (Phe9, Trp 13, His 29, Phe 34) on one face of the D0 domain ligated to loops comprising residues 14–18 and 88–92 of HLA-B\*5701 (Fig. 2a and Supplementary Table 3), both of which flexed slightly upon KIR3DL1\*001 binding (Supplementary Fig. 4)<sup>7</sup>. These two HLA loops exhibit very limited polymorphism among the HLA-A and HLA-B allotypes and mostly have main-chain interactions with the D0 domain, thereby indicating that the D0–HLA interactions are largely independent of sequence variation and likely to be conserved across most HLA allotypes. Lengthening or shortening the HLA-B\*5701 loop (residues 14–18) markedly reduced binding to KIR3DL1\*001 (Fig. 3a). Alanine substitution of Ser 11, His 29 and particularly Phe 9 in KIR3DL1\*001 impaired binding of HLA-B\*5701 tetramers, further highlighting the importance of the D0 contacts (Fig. 3b). Interestingly, the site of the D0-mediated interaction on HLA-B\*5701 has not, to the best of our knowledge, been observed in any HLA-binding immune receptor/co-receptor to date, indicating a unique molecular recognition signature, in which the D0 domain acts as an ‘innate sensor’ of an essentially invariant region of the HLA molecule.

The D1–D2 domains converged to form a continuous binding interface with HLA-B\*5701 (Fig. 1c, d), interacting with residues from the  $\alpha 1$ - and  $\alpha 2$ -helices flanking the P8 position of the peptide. The ligand-binding site of the D1–D2 domains was relatively flat, facilitating the close positioning of HLA-B\*5701, resulting in an intricate network of

interactions across the interface; as such, the total BSA upon complexation at the D1–D2 interaction site was quite large (total BSA, 1,360 Å<sup>2</sup>). The D1 and D2 domains contributed 600 and 760 Å<sup>2</sup> total BSA to the interface respectively, with the D1 domain docked above the  $\alpha 1$ -helix and contacting the peptide, whereas the D2 domain sat above the  $\alpha 2$ -helix, thereby providing immediate insight into the disparate roles that the D1 and D2 domains have in HLA-B\*5701 engagement (Fig. 1c, d). The D2 domain predominantly interacted with a region spanning residues 142–151 of HLA-B\*5701 (Supplementary Table 3), a region that shows limited polymorphism among HLA-B allotypes. At the core of the D2–HLA-B\*5701 binding interface, two aromatic residues of KIR3DL1\*001, Tyr 200 and Phe 276, converged onto the  $\alpha 2$ -helix, whereas polar interactions were located at the periphery (Fig. 2b). A feature of this interface was the centrally located Glu 282 of KIR3DL1\*001, a charged residue that abuts Leu 166 from the D1 domain, yet is stabilized by polar interactions with Tyr 200 and Ser 279 of KIR3DL1\*001, Lys 146 of HLA-B\*5701 and water-mediated interactions with the peptide and Arg 83 (not shown) on the  $\alpha 1$ -helix (Fig. 2c). Alanine substitution of Glu 201, Ser 227, Asp 230 or His 278, residues that were located at the exterior of the interface, had little effect on binding (Fig. 3b). In contrast, alanine substitution of Tyr 200 or Phe 276, which formed the central aromatic cluster, or the charged residue Glu 282, abrogated tetramer binding (Fig. 3b). Further, of the five HLA-B\*5701 mutations made at the D2–HLA-B\*5701 interface, three residues (Ile 142, Lys 146 and Ala 149) markedly affected the affinity of the interaction (Fig. 3a). These three HLA-B\*5701 residues interacted principally with Tyr 200 and Phe 276, further highlighting the importance of this internal core of KIR3DL1\*001 residues in driving the D2–HLA-B\*5701 interaction. Collectively, the D2–HLA-B\*5701-binding site seems to have co-evolved to form a highly complementary binding interface.

KIR3DL1 recognizes HLA class I allotypes that contain the Bw4 serological epitope spanning residues 77–83 on the  $\alpha 1$ -helix<sup>8,9</sup>.



**Figure 3 | Mutational analysis at the KIR3DL1\*001–pHLA-B\*5701 interface.** **a**, Surface plasmon resonance (SPR)-based measurements of the KIR3DL1\*001–HLA-B\*5701 mutants interaction. Results are expressed as percentage of the wild-type interaction; mutants are colour-coded according to the KIR3DL1\*001 domain they contact to correspond with Fig. 1. **b**, Capacity of HLA-B\*5701 tetramers to bind 293T cells expressing wild-type or mutant KIR3DL1\*001. HLA-B\*5701 tetramers, but not HLA-B\*0801 tetramers (data not shown), bound 293T cells transfected with KIR3DL1\*001. Binding is expressed as a proportion of positive cells relative to cells transfected with wild-type KIR3DL1\*001. Mutated residues are colour-coded as in Fig. 1.  $N = 2$  independent experiments; error bars represent s.e.m.

While the D1 domain was positioned over the Bw4 epitope making contacts with residues 79, 80 and 83, it interacted with a broader region of the  $\alpha$ 1-helix, including Gln 72, which bound to Met 165 (Fig. 2d and Supplementary Table 3). In marked contrast to the D2-mediated contacts, the D1–HLA-B\*5701 interface appeared to largely lack both charge and shape complementarity (Fig. 2d and Supplementary Fig. 5). Among the residues within the Bw4 motif, Arg 79 formed van der Waals contacts with Ser 140 and hydrogen-bonded to the main chain of Gly 138. Nevertheless the environment of Arg 79 was suboptimal, with its side chain being in close proximity to Lys 136 and Ile 139 of KIR3DL1\*001. Ile 80, a residue previously associated with KIR3DL1 reactivity<sup>10</sup>, formed a single van der Waals contact with Leu 166 and was positioned within a small hydrophobic cavity created by Glu 76, Arg 79 and Arg 83, a triad of HLA-B\*5701 residues that leaned towards each other to form an array of salt-bridging interactions (Fig. 2d). Further, Arg 83 from the Bw4 motif packed against and hydrogen-bonded to the main chain of His 278 (Fig. 2d). Surprisingly, none of the five alanine mutations introduced into the D1 domain had a substantial effect on the KIR3DL1–pHLA-B\*5701 interaction (Fig. 3b). However, in contrast, mutation of the corresponding HLA-B\*5701 contact residues did affect recognition, particularly the Ile80Ala and Arg83Ala mutations (Fig. 3a). Interestingly, mutation of Ile 80 to Thr, a natural dimorphism within the Bw4 motif, resulted in a modest reduction in the affinity of the interaction with KIR3DL1\*001 (Fig. 3a). Presumably, the Ile80Ala and Ile80Thr mutations differentially disrupt the conformation of the Glu 76–Arg 79–Arg 83 triad, thereby affecting KIR3DL1 recognition. Thus, whereas KIR3DL1\*001 contacted the highly polymorphic region of the HLA class I in a non-optimal manner, and the D1 residues were shown to be non-essential for this interaction, modifications within the HLA itself affected the D1–HLA-B\*5701 interaction and thus could serve to fine-tune the specificity of the interaction. Indeed, although KIR3DL1\*001 specifically binds HLA molecules that possess the Bw4 motif, it does not interact with the closely related Bw6 motif, which possesses a Gly at position 83. Accordingly our data provide a basis for understanding the importance of polymorphism at residue 83 for KIR3DL1 recognition of the Bw4<sup>+</sup> epitope<sup>2</sup>.

The D1 domain interacted with the LSSPVTKSF peptide; however, the sole direct interaction between the peptide and KIR3DL1\*001 was a van der Waals contact between P8–Ser (where P8 is position 8 of the peptide) and Leu 166 (Fig. 2c). Thus, KIR3DL1\*001 made limited contact with the peptide, analogous to the interactions observed between KIR2D and peptides bound to HLA-C<sup>4,5</sup>, and in marked contrast to CD94–NKG2A recognition of HLA-E<sup>11</sup>. To probe the role of peptide in the interaction, a series of peptides that were substituted at P8 were refolded with HLA-B\*5701 and assessed for their impact on recognition by KIR3DL1\*001. The Phe, His and Arg P8 substitutions all facilitated an interaction with KIR3DL1\*001, albeit with lower affinities, suggesting that the receptor interface has some capacity to tolerate large side chains at P8, consistent with the presence of a solvent-filled cavity adjacent to the P8 position at the KIR3DL1\*001–pHLA-B\*5701 interface. In contrast, the Ala, Glu and Leu P8 substitutions markedly reduced the corresponding interaction affinities (Supplementary Table 1 and Supplementary Fig. 1), suggesting that the KIR3DL1\*001 receptor can ‘discriminate’ between peptides. The basis for the differential effects of the P8 residue could be attributable either to direct steric hindrance/lack of complementarity between the peptide and KIR3DL1\*001, or to conformational alteration of the residues within the Bw4 motif itself<sup>12</sup>. Collectively, our observations are consistent with previous studies<sup>13,14</sup>, which demonstrated that the sequence of the bound peptide could have a profound effect on HLA recognition by KIR.

Next, we assessed the underlying HLA specificities of the KIR2DL and KIR3DL receptor families<sup>4,5</sup> (Supplementary Figs 6 and 7). The D1–D2 domains of KIR3DL1\*001 share clear sequence and structural homology with the HLA-C-reactive receptors, KIR2DL1, -2 and -3

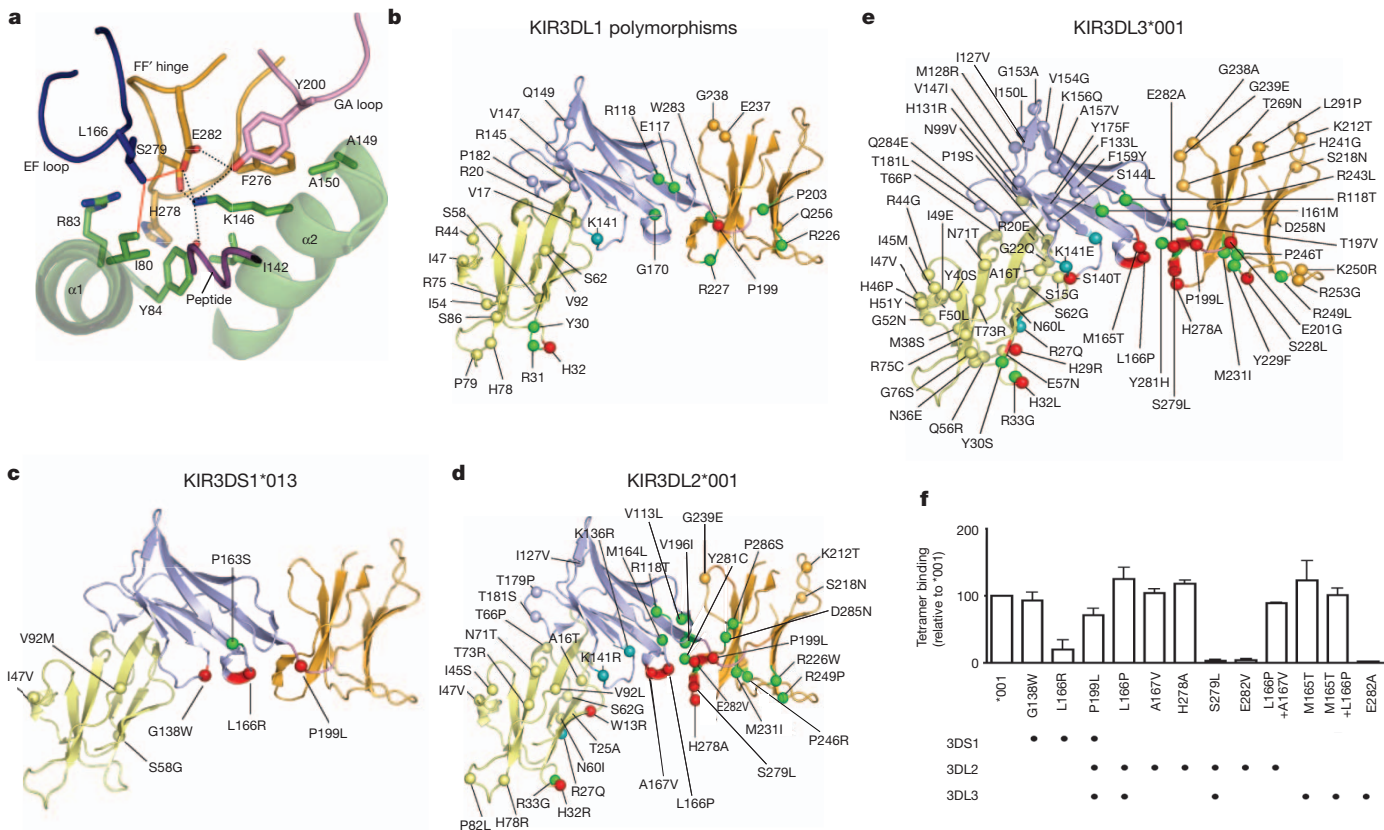
(refs 4, 5, 15), and there are a number of similarities in the recognition of the  $\alpha$ 2-helix by both KIR2DL1 and -2 and KIR3DL1\*001 (Supplementary Fig. 7). In contrast, the interactions between the KIR3DL1\*001 and KIR2DL1 receptors and the  $\alpha$ 1-helices of their respective HLA class I ligands vary (Supplementary Fig. 6b). These differences principally arise from the loop regions that connect the C and C'  $\beta$ -strands and the E and F  $\beta$ -strands in the D1 domain and the F and F'  $\beta$ -strands that bridge the D1 and D2 domains. For example, the CC' loop in the KIR2DL receptors adopts a notably different conformation from that observed in KIR3DL1\*001 (Supplementary Fig. 6c). In KIR3DL1\*001, this loop (137–140) is mostly flat and featureless, sitting adjacent to the  $\alpha$ 1-helical axis, forming limited contacts with HLA-B\*5701. In the KIR2DL receptors, the corresponding loop region (42–45) is orientated towards the  $\alpha$ 1-helix, and contains two prominent residues that would prevent binding to HLA-B\*5701 owing to steric hindrance with residues within the Bw4 motif. Thus, the D1-mediated contacts are critical for the HLA specificity differences between the KIR2DL1 family and KIR3DL1\*001.

The KIR3D family comprises the KIR3DL1/S1, KIR3DL2 and KIR3DL3 proteins<sup>16</sup>. More than 200 alleles within the KIR3D family have been described, with KIR3D allomorphs generally differing from each other by a limited number of amino acids<sup>17</sup>. Given the high sequence identity between KIR3DL1\*001 and KIR3DL2, KIR3DL3 and KIR3DS1 receptors (86, 74 and 97%, respectively), the KIR3DL1\*001–HLA-B\*5701 structure provided a template to examine the impact of sequence variation across the entire KIR3D family and relate this to pHLA specificity. Sequence and structural analyses suggested that a ‘hotspot’ resided within the D1–D2 domains, comprising loops 165–167, 199–201 and 278–282, all of which converged to form an intricate bonded network that centred on Glu 282 (Fig. 4a–e). Variation within these three loops could potentially alter the conformation of neighbouring residues within this hotspot region, thereby affecting receptor specificity.

KIR3DS1 is distinct among the KIR3D family in that it is an activating receptor. Genetic data have shown that *KIR3DL1* and *KIR3DS1* are allelic variants of the same gene and suggested that KIR3DS1 interacts with HLA-Bw4 molecules bearing an Ile at residue 80 (Bw4+I80)<sup>18</sup>. However, direct evidence of an interaction between KIR3DS1 and Bw4+I80 molecules is lacking<sup>19</sup>. Four positions that differ between KIR3DL1 and KIR3DS1 map to the KIR3DL1\*001–pHLA-B\*5701 interface and thus may affect the interaction (Fig. 4c), consistent with recent observations using HLA-A24 tetramers<sup>20,21</sup>. Whereas the Gly138Trp and Pro199Leu mutations had little impact on HLA-B\*5701 binding, mutation of Leu 166—which is located within the hotspot—to Arg substantially diminished tetramer binding (Fig. 4f), thereby providing a basis for why KIR3DS1 cannot bind HLA-B\*5701.

The KIR3DL2 family recognizes a limited subset of HLA-A allotypes<sup>22,23</sup>, with seven sequence differences that map to the hotspot region (Fig. 4d). The introduction of these residues into KIR3DL1\*001 showed that whereas the Leu166Pro, Ala167Val (Fig. 4f) and His278Ala mutations (Fig. 3b) did not impair recognition of HLA-B\*5701, the Ser279Leu and Glu282Val mutations markedly reduced tetramer binding (Fig. 4f). Removal of the charged moiety of Glu 282 would disrupt the intricate network of interactions at the KIR3DL1\*001–pHLA-B\*5701 interface. Whereas the Ser279Ala mutation did not abrogate HLA-B\*5701 binding (Fig. 3b), the impact of the Ser279Leu mutation was much more pronounced. This effect appears attributable to the more bulky Leu residue causing a steric clash with Arg 83, thereby suggesting a basis for the lack of reactivity of KIR3DL2 towards the Bw4 motif. Moreover, unlike HLA-B\*5701 and other HLA-Bw4 allotypes, HLA-A3 and HLA-A11 possess a Gly at position 83 rather than Arg, which is a crucial determinant for KIR3DL1 recognition of the Bw4 motif<sup>2</sup>.

The specificity of the KIR3DL3 receptor family is undefined, and a number of differences between KIR3DL1 and -3DL3 reside within the hotspot region (Fig. 4e). Binding experiments showed that the



**Figure 4 | Mapping of polymorphisms and sequence variations onto the structure of KIR3DL1\*001.** a, The 'hotspot' comprises three loops. b, Polymorphisms within the KIR3DL1 family. c–e, Differences between KIR3DS1\*013 and KIR3DL1\*001 (c), KIR3DL2\*001 and KIR3DL1\*001 (d), KIR3DL3\*001 and KIR3DL1\*001 (e). Polymorphisms are represented as spheres: red, direct contacts; cyan, water-mediated contacts; green, residues

Met165Thr or Leu166Pro substitutions in KIR3DL1\*001 did not affect HLA-B\*5701 binding (Fig. 4f). Further, whereas the Pro199Leu substitution had a modest impact on recognition, the Glu282Ala substitution within KIR3DL1\*001 totally abrogated tetramer binding (Fig. 3b), thereby indicating that residues 279 and 282 are critical determinants of the specificity differences between KIR3DL1 and other KIR3D family members.

Surprisingly, the extensive polymorphism among the inhibitory receptors within each KIR3D family was predominantly located at sites not directly implicated in pHLA binding (Fig. 4b). Collectively, these observations indicate that the majority of KIR3D polymorphisms within a family<sup>24,25</sup> are unlikely to directly affect the affinity of the pHLA interaction per se, but rather are likely to affect pHLA binding via altering expression levels and/or the clustering of the KIR3D receptors on the cell surface, whereas sequence differences across the KIR3D family directly affect pHLA affinity and specificity. Indeed, functional studies have shown that polymorphisms in residues such as 238 that are distant from the receptor/ligand interface can affect target cell recognition by KIR3DL1<sup>+</sup> NK cells<sup>26</sup>.

Collectively, our data provide a fundamental basis for understanding how a representative KIR3DL family member interacts with an HLA-B molecule that possesses the Bw4 motif. We show that the D0 domain, a feature of this family, interacts with a previously unrecognized determinant on the HLA molecule, which is highly conserved across HLA-A and HLA-B allotypes in particular. These observations indicate that the D0 domain acts as an innate HLA sensor at a site that that is not involved in either peptide or TCR binding<sup>27</sup>. The KIR3DL interaction sites seem to be largely conserved across the KIR3D family, with specificity differences mapping to a hotspot

that may affect binding; other, remaining residues coloured according to domain. **f**, The capacity of HLA-B\*5701 tetramers to bind to 293T cells transfected with plasmids encoding either a Flag-tagged KIR3DL1\*001 or 10 site-directed mutants representing sites of 3DL1/2/3/3DS1 variation that contacted HLA-B\*5701.  $N = 2$  independent experiments; error bars represent s.e.m. Variations across the KIR3D family are shown underneath.

within the interaction interface. In contrast, the polymorphisms within individual *KIR3D* gene families are largely at positions that are spatially separate from the binding site, a number of which are the subject of positive selection<sup>17</sup>. This suggests that other evolutionary pressures, such as pathogen-mediated immune evasion strategies, may drive *KIR3D* diversification at sites distant from the ligand-binding site.

## METHODS SUMMARY

**Protein expression and purification.** Inclusion body preparations of the HLA-B\*5701 heavy chain and  $\beta_2$ -microglobulin were refolded and purified as detailed previously<sup>7</sup>. Residues 1–299 of KIR3DL1\*001 were cloned into the pHLSec mammalian expression vector<sup>28</sup> with N-terminal 6 $\times$ His and secretion tags. KIR3DL1\*001 was expressed from transiently transfected HEK 293S cells. Purified KIR3DL1\*001 was then concentrated to 15 mg ml<sup>-1</sup> and deglycosylated with endoglycosidase H (New England Biolabs).

**Crystallization and data collection.** The KIR3DL1\*001-pHLA-B\*5701 complex was crystallized and its structure determined. Further details are provided in Methods.

**Transfection studies.** The sequence for a Flag tag (GACTACAAAGACGATGACGACAAG) was added to the 5' end of KIR3DL1\*001 by primer addition and this cDNA was then cloned into a pEF6 vector. Specific nucleotide residues were mutated using the QuikChange II Site Directed Mutagenesis Kit (Stratagene). Plasmids were transfected into HEK293T cells using the FuGene 6 transfection reagent (Roche) according to the manufacturer's instructions. After 48 h, the cells were harvested and stained with anti-Flag (clone M2, Sigma Aldrich) antibody or with tetramer for 30 min at 4 °C. The cells were then washed and analysed on a Fortessa flow cytometer (BD Biosciences).

**Surface plasmon resonance.** Surface plasmon resonance experiments were conducted at 25 °C on a Biacore 3000 instrument using HBS buffer (10 mM HEPES-HCl (pH 7.4), 150 mM NaCl and 0.005% surfactant P20 supplied by the manufacturer). Further details are provided in Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 6 June; accepted 26 August 2011.**

**Published online 23 October 2011.**

1. Parham, P. MHC class I molecules and KIRs in human history, health and survival. *Nature Rev. Immunol.* **5**, 201–214 (2005).
2. Sanjanwala, B., Draghi, M., Norman, P. J., Guethlein, L. A. & Parham, P. Polymorphic sites away from the Bw4 epitope that affect interaction of Bw4<sup>+</sup> HLA-B with KIR3DL1. *J. Immunol.* **181**, 6293–6300 (2008).
3. Martin, M. P. *et al.* Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. *Nature Genet.* **39**, 733–740 (2007).
4. Boyington, J. C., Motyka, S. A., Schuck, P., Brooks, A. G. & Sun, P. D. Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. *Nature* **405**, 537–543 (2000).
5. Fan, Q. R., Long, E. O. & Wiley, D. C. Crystal structure of the human natural killer cell inhibitory receptor KIR2DL1–HLA–Cw4 complex. *Nature Immunol.* **2**, 452–460 (2001).
6. Colonna, M. & Samaridis, J. Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. *Science* **268**, 405–408 (1995).
7. Chessman, D. *et al.* Human leukocyte antigen class I-restricted activation of CD8<sup>+</sup> T cells provides the immunogenetic basis of a systemic drug hypersensitivity. *Immunity* **28**, 822–832 (2008).
8. Litwin, V., Gumperz, J., Parham, P., Phillips, J. H. & Lanier, L. L. NKB1: a natural killer cell receptor involved in the recognition of polymorphic HLA-B molecules. *J. Exp. Med.* **180**, 537–543 (1994).
9. Gumperz, J. E., Litwin, V., Phillips, J. H., Lanier, L. L. & Parham, P. The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor. *J. Exp. Med.* **181**, 1133–1144 (1995).
10. Cella, M., Longo, A., Ferrara, G. B., Strominger, J. L. & Colonna, M. NK3-specific natural killer cells are selectively inhibited by Bw4-positive HLA alleles with isoleucine 80. *J. Exp. Med.* **180**, 1235–1242 (1994).
11. Petrie, E. J. *et al.* CD94–NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. *J. Exp. Med.* **205**, 725–735 (2008).
12. Hülsmeier, M. *et al.* Thermodynamic and structural equivalence of two HLA-B27 subtypes complexed with a self-peptide. *J. Mol. Biol.* **346**, 1367–1379 (2005).
13. Peruzzi, M., Parker, K. C., Long, E. O. & Malnati, M. S. Peptide sequence requirements for the recognition of HLA-B\*2705 by specific natural killer cells. *J. Immunol.* **157**, 3350–3356 (1996).
14. Fadda, L. *et al.* Peptide antagonism as a mechanism for NK cell activation. *Proc. Natl Acad. Sci. USA* **107**, 10160–10165 (2010).
15. Maenaka, K. *et al.* Killer cell immunoglobulin receptors and T cell receptors bind peptide-major histocompatibility complex class I with distinct thermodynamic and kinetic properties. *J. Biol. Chem.* **274**, 28329–28334 (1999).
16. Trowsdale, J. *et al.* The genomic context of natural killer receptor extended gene families. *Immunol. Rev.* **181**, 20–38 (2001).
17. Norman, P. J. *et al.* Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. *Nature Genet.* **39**, 1092–1099 (2007).
18. Martin, M. P. *et al.* Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nature Genet.* **31**, 429–434 (2002).
19. Carr, W. H. *et al.* Cutting edge: KIR3DS1, a gene implicated in resistance to progression to AIDS, encodes a DAP12-associated receptor expressed on NK cells that triggers NK cell activation. *J. Immunol.* **178**, 647–651 (2007).
20. Sharma, D. *et al.* Dimorphic motifs in D0 and D1+D2 domains of killer cell Ig-like receptor 3DL1 combine to form receptors with high, moderate, and no avidity for the complex of a peptide derived from HIV and HLA-A\*2402. *J. Immunol.* **183**, 4569–4582 (2009).
21. O'Connor, G. M. *et al.* Analysis of binding of KIR3DS1\*014 to HLA suggests distinct evolutionary history of KIR3DS1. *J. Immunol.* **187**, 2162–2171 (2011).
22. Dohring, C., Scheidegger, D., Samaridis, J., Cella, M. & Colonna, M. A human killer inhibitory receptor specific for HLA-A1.2. *J. Immunol.* **156**, 3098–3101 (1996).
23. Hansasuta, P. *et al.* Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. *Eur. J. Immunol.* **34**, 1673–1679 (2004).
24. Yawata, M. *et al.* Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. *J. Exp. Med.* **203**, 633–645 (2006).
25. Khakoo, S. I., Geller, R., Shin, S., Jenkins, J. A. & Parham, P. The D0 domain of KIR3D acts as a major histocompatibility complex class I binding enhancer. *J. Exp. Med.* **196**, 911–921 (2002).
26. Carr, W. H., Pando, M. J. & Parham, P. KIR3DL1 polymorphisms that affect NK cell inhibition by HLA-Bw4 ligand. *J. Immunol.* **175**, 5222–5229 (2005).
27. Godfrey, D. I., Rossjohn, J. & McCluskey, J. The fidelity, occasional promiscuity, and versatility of T cell receptor recognition. *Immunity* **28**, 304–314 (2008).
28. Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D* **62**, 1243–1250 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the staff at the MX2 beamline of the Australian synchrotron for assistance with data collection. We thank A. Radu Aricescu for the gift of the pHlsec vector. This research was supported by the National Health and Medical Research Council of Australia (NHMRC), the Australian Research Council (ARC) and the Intramural Research Programs of the National Cancer Institute and the National Institute of Allergy and Infectious Diseases, National Institutes of Health. D.W.M. and G.M.O.C. were supported by the Intramural AIDS Targeted Antiviral Program of the National Institutes of Health. J.P.V. is supported by an NHMRC Peter Doherty Research Fellowship; D.A.P. is supported by a Medical Research Council (UK) Senior Clinical Fellowship; B.A.P.L. is supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health; C.S.C. is supported by an ARC QEII Fellowship; J.R. is supported by an ARC Federation Fellowship.

**Author Contributions** J.P.V. solved the structure, undertook analysis, performed experiments and contributed to manuscript preparation. H.H.R., T.B., R.C.D., R.B., P.M.S., M.A.O., J.M.L.W., C.M.H., J.L., S.M.M., S.G. and C.S.C. performed experiments and/or analysed data. G.M.O.C., D.A.P., B.A.P.L. and D.W.M. performed experiments and/or analysed data and contributed to the writing of the manuscript; A.G.B. and J.R. were the joint senior authors—they co-led the investigation, devised the project, analysed the data and wrote the manuscript.

**Author Information** The atomic coordinates and structure factors for the KIR3DL\*001–pHLA-B\*5701 complex were deposited in the Protein Data Bank under accession code 3VH8. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to A.G.B. ([agbrooks@unimelb.edu.au](mailto:agbrooks@unimelb.edu.au)) or J.R. ([jamie.rossjohn@monash.edu](mailto:jamie.rossjohn@monash.edu)).

## METHODS

**Protein expression and purification.** HLA-B\*5701 and  $\beta_2$ -microglobulin were expressed separately in *E. coli* from the pET-30 vector. Inclusion body preparations of the HLA-B\*5701 and  $\beta_2$ -microglobulin were refolded and purified as detailed previously<sup>7</sup>. In brief, the resultant HLA class I complexes were purified by DEAE sepharose (Sigma) anion exchange chromatography using 10 mM Tris pH 8.0 and eluted with 150 mM NaCl. The protein was then further purified by gel filtration using an S200 16/60 column (GE Healthcare). The final purification step used anion exchange chromatography on a MonoQ column (GE Healthcare). The binary complex was concentrated in 10 mM Tris pH 8.0, 150 mM NaCl for use in crystallization trials and surface plasmon resonance (SPR) studies. The mutants of HLA-B\*5701 were generated using the QuikChange PCR method (Stratagene) and purified as described earlier.

Residues 1–299 of KIR3DL1\*001 were cloned into the pHLsec mammalian expression vector with N-terminal 6×His and secretion tags. KIR3DL1\*001 was expressed from transiently transfected HEK 293S cells. Secreted KIR3DL1\*001 was harvested from the culture media 3 days after transfection by first dialysing the media against 10 mM Tris pH 8.0, 300 mM NaCl before the use of nickel affinity resin. The KIR3DL1\*001 was eluted from the nickel resin with 10 mM Tris pH 8.0, 300 mM NaCl, 50 mM EDTA. The protein was purified by gel filtration chromatography using an S200 16/60 column (GE Healthcare) in 10 mM Tris pH 8.0, 300 mM NaCl. Purified KIR3DL1\*001 was then concentrated to 15 mg ml<sup>−1</sup> and deglycosylated with endoglycosidase H (New England Biolabs). The extent of deglycosylation was monitored by SDS–PAGE and this material was used in crystallization trials. For SPR studies a similar construct of KIR3DL1\*001 was prepared in the pFastBac vector and expressed from Hi-5 insect cells (Invitrogen). The KIR3DL1\*001 was purified as described earlier with the exception that the endoglycosidase H deglycosylation step was not performed.

**Crystallization and data collection.** The KIR3DL1\*001–pHLA-B\*5701 complex at 15 mg ml<sup>−1</sup> was crystallized at 294 K by the hanging-drop vapour-diffusion method from a solution comprising 16% PEG 3350, 2% tacsimate pH 5 and 0.1 M tri-sodium citrate pH 5.6. The crystals typically grew to dimensions 0.3 × 0.3 × 0.2 mm in 7 days. Before data collection, the crystals were equilibrated in crystallization solution with 35% PEG 3350 added as a cryoprotectant and then flash-cooled in a stream of liquid nitrogen at 100 K. X-ray diffraction data were recorded on a Quantum-315 CCD detector at the MX2 beamline of the Australian Synchrotron. The data were integrated and scaled using DENZO and SCALEPACK from the HKL2000 program suite. Details of the data processing statistics are given in Supplementary Table 2.

The final model comprises residues 6–261, 267–292 and there are three glycosylation sites located at Asn 71, Asn 158 and Asn 252.

**Structure determination and refinement.** The structure was determined by molecular replacement using MOLREP. The search models used were the structures of HLA-B\*5701 and KIR2DL1 (PDB codes 2RFX and 1IM9). The positions of the two complexes in the asymmetric unit were found in an incremental manner.

The orientation of the first HLA molecule was found and subsequently the position of the D1 and D2 domains of the KIR receptor were placed. The second complex was fitted by application of the pseudo-translation vector 0.0, 0.5, 0.5.

Refinement of the model was carried out in REFMAC with strict twofold non-crystallographic symmetry (NCS) applied. Structure building proceeded with iterative rounds of manual building in COOT and refinement in REFMAC. The D0 domain of KIR3DL1\*001 was manually built from the resultant electron density maps. The NCS restraints were removed for the final rounds of refinement. Solvent was added with COOT and the structure validated with MOLPROBITY<sup>29</sup>. The final structure comprises two KIR3DL1\*001–pHLA-B\*5701 complexes in the asymmetric unit, the association of which did not indicate higher-order oligomeric assemblies within the crystal lattice. The final refinement values are summarized in Supplementary Table 2. The crystals contained two virtually indistinguishable ternary complexes within the asymmetric unit, so structural analyses were confined to one KIR3DL1\*001–pHLA-B\*5701 complex.

**Transfection studies.** The sequence for a Flag tag (GACTACAAAGACGATGACGACAAG) was added to the 5′ end of KIR3DL1\*001 by primer addition and this cDNA was then cloned into a pEF6 vector. Specific nucleotide residues were mutated using the QuikChange II Site Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions using PAGE-purified primers. Sequences were verified by direct sequencing. These constructs were introduced into HEK293T cells using FuGene 6 transfection reagent (Roche) according to the manufacturer's instructions. After 48 h, the cells were harvested and stained with anti-Flag (clone M2, Sigma Aldrich) antibody or with tetramer for 30 min at 4 °C. The cells were then washed and analysed on a Fortessa flow cytometer (BD Biosciences). Analysis of cell surface expression as assessed by staining with anti-Flag monoclonal antibody showed that the introduction of the mutations had no substantial effect on expression (data not shown). All transfection data are representative of two independent experiments.

**SPR.** SPR experiments were conducted at 25 °C on a Biacore 3000 instrument using HBS buffer (10 mM HEPES-HCl (pH 7.4), 150 mM NaCl and 0.005% surfactant P20 supplied by the manufacturer). The HLA class I-specific antibody W6/32 was immobilized on a CM5 chip via amine coupling according to manufacturer's instructions. The pHLA complexes, and mutants thereof, were captured by W6/32 creating a surface density of approximately 500–1,000 resonance units. Various concentrations of KIR3DL1\*001 (2.37 to 300  $\mu$ M) were injected over the captured pHLA at 5  $\mu$ l min<sup>−1</sup>. The final response was calculated by subtracting the response of W6/32 alone from the KIR3DL1\*001–pHLA-B\*5701 complex. The equilibrium data were analysed using GraphPad Prism. The shortened form of HLA-B\*5701 comprised Gly-Gly-Gly in place of residues 14–19; in the long form of HLA-B\*5701, Gly-Gly-Gly was inserted after Gly 16. For the SPR experiments, data are representative of two independent experiments with error bars representing s.e.m. of the duplicates.

29. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

# Rad51 paralogues Rad55–Rad57 balance the antirecombinase Srs2 in Rad51 filament formation

Jie Liu<sup>1</sup>, Ludovic Renault<sup>2</sup>, Xavier Veaute<sup>3</sup>, Francis Fabre<sup>3</sup>, Henning Stahlberg<sup>2,4</sup> & Wolf-Dietrich Heyer<sup>1,2</sup>

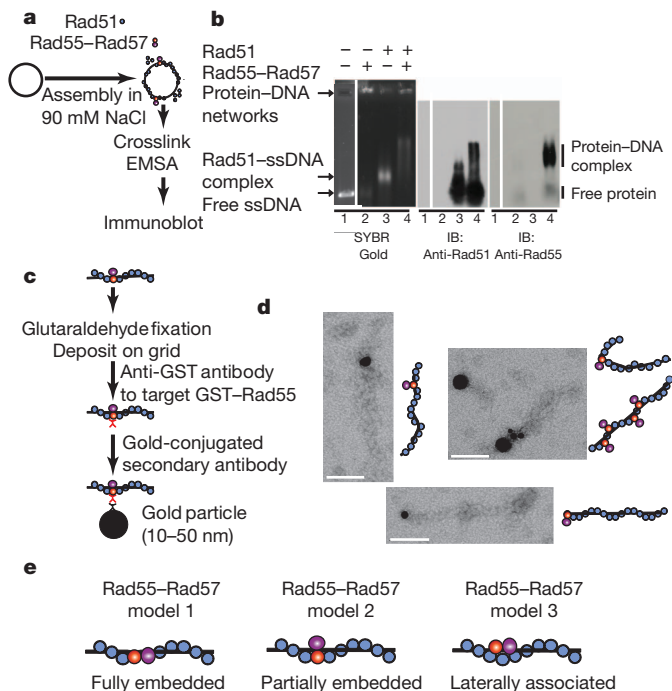
Homologous recombination is a high-fidelity DNA repair pathway. Besides a critical role in accurate chromosome segregation during meiosis, recombination functions in DNA repair and in the recovery of stalled or broken replication forks to ensure genomic stability. In contrast, inappropriate recombination contributes to genomic instability, leading to loss of heterozygosity, chromosome rearrangements and cell death. The RecA/UvsX/RadA/Rad51 family of proteins catalyses the signature reactions of recombination, homology search and DNA strand invasion<sup>1,2</sup>. Eukaryotes also possess Rad51 paralogues, whose exact role in recombination remains to be defined<sup>3</sup>. Here we show that the *Saccharomyces cerevisiae* Rad51 paralogues, the Rad55–Rad57 heterodimer, counteract the antirecombination activity of the Srs2 helicase. The Rad55–Rad57 heterodimer associates with the Rad51–single-stranded DNA filament, rendering it more stable than a nucleoprotein filament containing Rad51 alone. The Rad51–Rad55–Rad57 co-filament resists disruption by the Srs2 antirecombinase by blocking Srs2 translocation, involving a direct protein interaction between Rad55–Rad57 and Srs2. Our results demonstrate an unexpected role of the Rad51 paralogues in stabilizing the Rad51 filament against a biologically important antagonist, the Srs2 antirecombination helicase. The biological significance of this mechanism is indicated by a complete suppression of the ionizing radiation sensitivity of *rad55* or *rad57* mutants by concomitant deletion of *SRS2*, as expected for biological antagonists. We propose that the Rad51 presynaptic filament is a meta-stable reversible intermediate, whose assembly and disassembly is governed by the balance between Rad55–Rad57 and Srs2, providing a key regulatory mechanism controlling the initiation of homologous recombination. These data provide a paradigm for the potential function of the human RAD51 paralogues, which are known to be involved in cancer predisposition and human disease.

Rad51 protein and its homologues RecA, UvsX and RadA form nucleoprotein filaments with ssDNA that perform homology search and DNA strand invasion during homologous recombination. The Rad51 paralogues share the RecA core with the Rad51 protein featuring unique amino- and carboxy-terminal extensions (Supplementary Fig. 2), but themselves do not form filaments and are unable to perform homology search and DNA strand invasion<sup>2–4</sup>. Whereas humans contain five paralogues (RAD51B, RAD51C, RAD51D, XRCC2, XRCC3), the budding yeast *Saccharomyces cerevisiae* contains two clearly identifiable paralogues, Rad55 and Rad57 (Supplementary Fig. 2). Rad55 and Rad57 in yeast as well as the five human RAD51 paralogues have unique non-redundant functions in recombination, and mutations in any one of them lead to recombination defects, chromosomal instability, sensitivity to DNA damage, and meiotic defects<sup>1–3</sup>. Defects in the budding yeast *RAD55* and *RAD57* genes lead to identical and epistatic phenotypes in DNA repair and recombination, consistent with the formation of a stable Rad55–Rad57 heterodimer<sup>4,5</sup>. Rad55–Rad57 heterodimers were inferred to function as

mediator proteins<sup>6</sup> allowing assembly of the Rad51 nucleoprotein filament on ssDNA covered by the eukaryotic ssDNA-binding protein RPA<sup>4</sup>. This suggested that Rad55–Rad57 are involved in the nucleation of the Rad51 filament, which is otherwise inhibited on RPA-covered ssDNA. This nucleation model is akin to the role of RecFOR or BRCA2 in nucleating RecA or human RAD51 filaments<sup>7–9</sup>. Rad51 filament formation *in vivo* can be monitored cytologically as Rad51 focus formation at the site of DNA damage<sup>10</sup>. Unexpectedly, Rad51 focus formation after ionizing radiation in yeast was demonstrated to be independent of Rad55–Rad57 and formation of visible Rad55–Rad57 foci required Rad51 (ref. 10). These results are difficult to reconcile with the nucleation model derived from the biochemical results and suggest an alternative function of Rad55–Rad57 *in vivo*.

To address the function of the Rad51 paralogues in yeast, we determined the effect of Rad55–Rad57 on the stability of Rad51–ssDNA nucleoprotein complexes. Deletion mutants of the *RAD55* or *RAD57* genes display a curious enhancement of some phenotypes at low temperature (in particular ionizing radiation sensitivity; see Supplementary Fig. 12)<sup>5</sup>, indicating that these proteins are involved in the stabilization of a molecular complex, probably the Rad51 presynaptic filament. To test this hypothesis, we incubated subsaturating amounts of Rad51 protein with ssDNA (1 Rad51 per 15 nucleotides) in the presence of substoichiometric amounts of Rad55–Rad57 heterodimer (1 Rad55–Rad57 per 4 Rad51) and challenged the filaments with buffer containing a high salt concentration (500 mM NaCl) (Supplementary Fig. 3a, b). Under these conditions, Rad51 does not maintain stable complexes with ssDNA during electrophoresis. However, the presence of Rad55–Rad57 resulted in stable, Rad51-containing ssDNA complexes that withstood the salt challenge. In a complementary approach, we examined the effect of Rad55–Rad57 on Rad51 filament formation at near-physiological ionic strength (90 mM NaCl) (Fig. 1a, b). Under these conditions, only a fraction of the available Rad51 binds ssDNA, causing retarded mobility of the DNA (Fig. 1b, lane 3). Addition of substoichiometric amounts of Rad55–Rad57 (1 Rad55–Rad57 per 6 Rad51 in lane 4 of Fig. 1b) led to the formation of a novel, supershifted complex that contained both Rad51 and Rad55–Rad57, as demonstrated by immunoblotting. Rad55–Rad57 heterodimer alone binds to DNA under these conditions, leading to the formation of protein networks that are too large to enter the gel (Fig. 1b, lane 2). The results from both experiments (Fig. 1b; Supplementary Fig. 3) indicate that Rad55–Rad57 form a co-complex with Rad51 on ssDNA and stabilize Rad51–ssDNA filaments. Indeed, immunogold electron microscopy targeted towards Rad55 (glutathione S-transferase (GST)-tag; see Fig. 1c) directly visualized Rad55 associated with the Rad51–ssDNA filaments (Fig. 1d). Control experiments demonstrated the specificity of the gold labelling (Supplementary Table 1) with over 90% of the gold particles associated with clearly identifiable Rad51 filaments. The remainder may have associated with filaments too short to be scored or with free Rad55–Rad57. Gold particles were found either at the filament terminus ( $n = 40$ ) or

<sup>1</sup>Department of Microbiology, University of California, Davis, Davis, California 95616-8665, USA. <sup>2</sup>Department of Molecular & Cellular Biology, University of California, Davis, Davis, California 95616-8665, USA. <sup>3</sup>CEA-DSV-Institut de Radiobiologie Cellulaire et Moléculaire, UMR217 CNRS/CEA, F-92265 Fontenay aux Roses, France. <sup>4</sup>Center for Cellular Imaging and Nanoanalytics, University Basel, CH-4056 Basel, Switzerland.

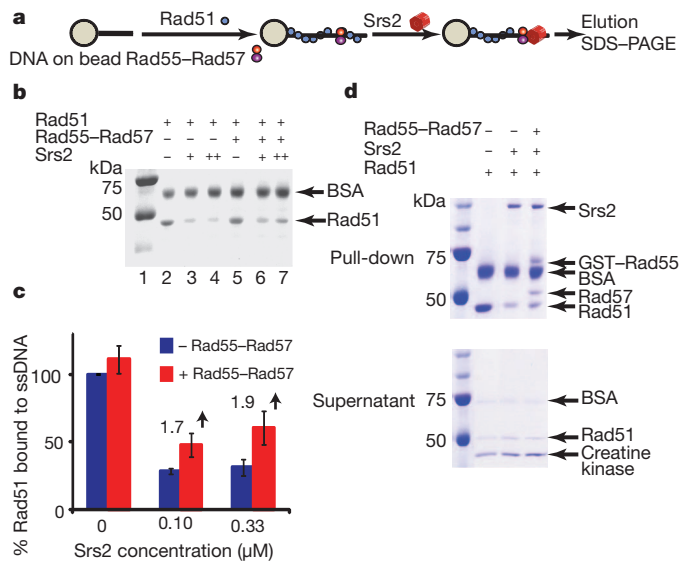


**Figure 1 | Rad55–Rad57 is associated with and stabilizes Rad51–ssDNA filaments.** **a**, Rad51–ssDNA filament assembly assay. EMSA, electrophoretic mobility shift assay. **b**, Rad51 (0.67  $\mu$ M) with or without Rad55–Rad57 (0.11  $\mu$ M) was incubated with 4  $\mu$ M  $\phi$ X174 ssDNA. The migration position of free protein was confirmed in controls lacking DNA (Supplementary Fig. 3c). IB, immunoblot. **c**, Reaction scheme of immunoaffinity gold labelling of Rad55. **d**, Electron microscopy images of gold-labelled Rad55 associated with Rad51–ssDNA filament (1:3 Rad51:nucleotide; 2.34  $\mu$ M Rad51  $\pm$  0.43  $\mu$ M Rad55–Rad57, 7  $\mu$ M ssDNA). Scale bars, 100 nm. **e**, Models for the disposition of Rad55–Rad57 with the Rad51 filament. For simplicity, only model 2 is drawn in further illustrations.

interstitially ( $n = 43$ ) (Supplementary Table 1). Negative controls with Rad51 filaments assembled in the absence of Rad55–Rad57 showed negligible gold labelling (Supplementary Table 1). These data show that Rad55–Rad57 are associated with the Rad51–ssDNA filament, but the exact disposition of the heterodimer with the filament remains to be determined (see Fig. 1e).

Salt stability of protein–DNA complexes is a valuable biochemical criterion. To establish biological significance, we tested whether Rad55–Rad57 heterodimers stabilize Rad51–ssDNA filaments against a biologically relevant destabilizer. The Srs2 helicase was identified as a negative regulator of homologous recombination, and genetic experiments indicated that Srs2 targets Rad51 protein<sup>11–13</sup>. Consistent with the genetic data, Srs2 translocates on ssDNA and disrupts Rad51 pre-synaptic filaments *in vitro*, providing a compelling mechanism for its function as an antirecombinase<sup>14–16</sup>. In the presence of 0.1 or 0.33  $\mu$ M Srs2 approximately 70% of the Rad51 is dissociated as assessed by measuring Rad51 associated with ssDNA coupled to magnetic beads (Fig. 2a–c). The presence of substoichiometric amounts of Rad55–Rad57 (0.1  $\mu$ M) enhanced the recovery of ssDNA-bound Rad51 by ~twofold (from 31% to 60% in the presence of 0.33  $\mu$ M Srs2). Rad55–Rad57 and Srs2 bound to Rad51-covered ssDNA in a quantitative and concomitant manner (Fig. 2d). Together the data show that Rad55–Rad57 inhibit Srs2 when bound to DNA and not in solution. Concentration-dependent inhibition of Srs2-mediated dissociation of Rad51 from ssDNA by Rad55–Rad57 was also observed in a topology-based assay (Supplementary Figs 4, 5).

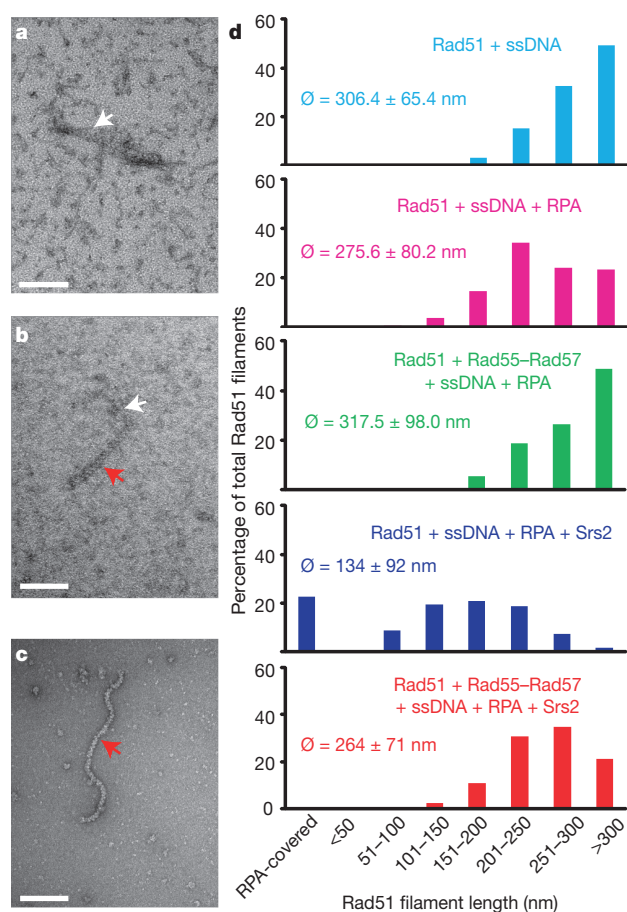
To investigate the role of Rad55–Rad57 in antagonizing disruption of Rad51 presynaptic filaments by Srs2 further, we used electron microscopy to examine nucleoprotein filaments directly (Fig. 3 and



**Figure 2 | Rad55–Rad57 stabilize Rad51–ssDNA filaments to resist disruption by Srs2.** **a**, Pull-down assay measuring stability of Rad51–ssDNA complexes (1:3 Rad51:nucleotide, 1  $\mu$ M Rad51  $\pm$  0.1  $\mu$ M Rad55–Rad57) against disruption by Srs2 (0.1 or 0.33  $\mu$ M). **b**, Rad51 remaining bound to ssDNA. **c**, Quantification of results in **b** and additional experiments. Shown are means  $\pm$  1 s.d.,  $n = 3$ . **d**, Concomitant binding of Rad55–Rad57 and Srs2 to Rad51-covered ssDNA. Pull-down assay measuring stability of Rad51–ssDNA complexes (1:3 Rad51:nucleotide, 1  $\mu$ M Rad51  $\pm$  0.2  $\mu$ M Rad55–Rad57) against disruption by 0.33  $\mu$ M Srs2. Top, pull-downs; bottom, supernatants.

Supplementary Fig. 6). Rad51 filaments were assembled on a 600-nucleotide fragment of ssDNA and RPA was added to visualize free ssDNA. Consistent with previous observations<sup>14,15</sup>, in the absence of Rad55–Rad57 Srs2 disrupts the Rad51–ssDNA filament efficiently, leading to binding of RPA to the newly exposed ssDNA (Fig. 3). Importantly, when substoichiometric amounts of Rad55–Rad57 were incubated with Rad51 and ssDNA, the filaments were stabilized against disruption by Srs2, as indicated by the significantly increased mean filament length.

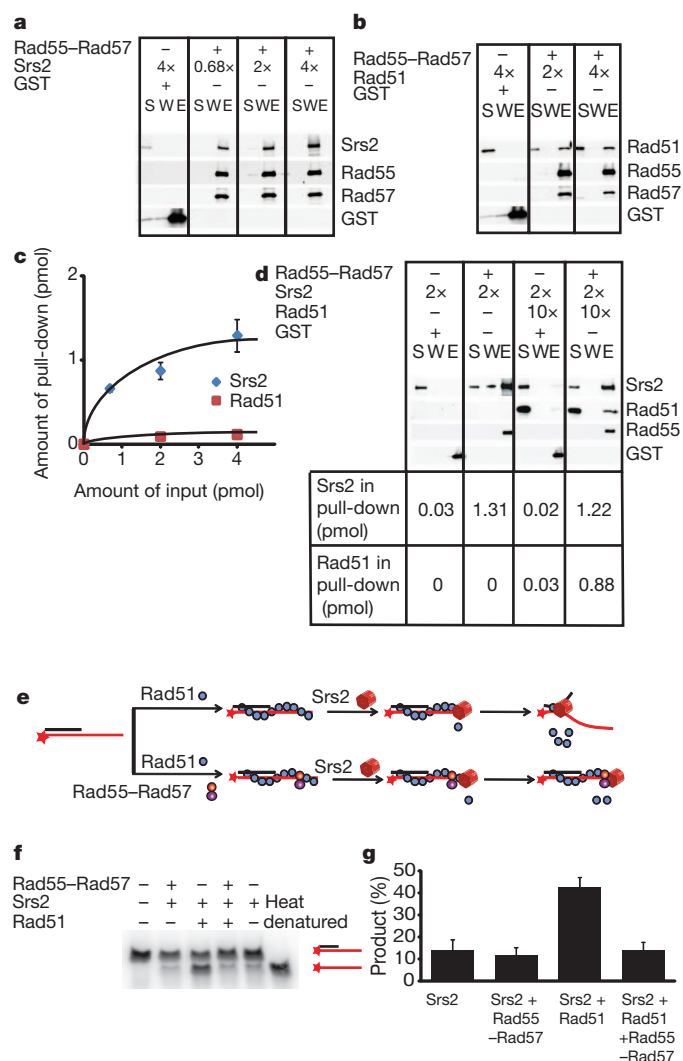
How do Rad55–Rad57 heterodimers block Srs2 from dissociating Rad51 from ssDNA? Srs2 is known to interact with Rad51 and trigger the Rad51 ATPase leading to dissociation of Rad51 from ssDNA<sup>16</sup>. We found that Rad55–Rad57 form a 1:1 complex with Srs2 (Fig. 4a) and have higher affinity to Srs2 than to Rad51 (Fig. 4b, c). Excess Rad51 does not compete with Srs2 binding to Rad55–Rad57 (Supplementary Fig. 7). Moreover, Rad55–Rad57 heterodimers are able to simultaneously bind Rad51 and Srs2 in a 1:1:1 stoichiometry (Fig. 4d and Supplementary Figs 7–9). We considered the possibility that Rad55–Rad57 inhibit the Srs2 ATPase activity and by that Srs2 translocation, but Srs2 ATPase activity is barely altered by the presence of Rad55–Rad57 (data not shown). Srs2 translocase/helicase activity is stimulated by Rad51 binding to DNA<sup>17</sup> (Fig. 4e–g). Importantly, Rad55–Rad57 completely suppress this stimulatory effect of Rad51, leading to inhibition of the Srs2 helicase activity even at a fivefold molar excess of Srs2 over Rad55–Rad57 (Fig. 4f, g and Supplementary Fig. 10). This substoichiometric action of Rad55–Rad57 eliminates the possibility that Rad55–Rad57 inhibition functions by binding Srs2 in solution. Rad55–Rad57 only slightly inhibit Srs2 helicase in the absence of Rad51 (Fig. 4g and Supplementary Fig. 10c). Control experiments show that this effect depends on Srs2 translocating in the expected 3' to 5' direction (Supplementary Fig. 10d), showing that Rad55–Rad57 inhibit Srs2 translocation on DNA to increase filament stability (Figs 1 and 2) and function (Supplementary Fig. 11). Direct visualization of human RAD51 filaments revealed that RAD51 is only able to form discontinuous short clusters on double-stranded DNA, as a result of frequent nucleation but limited extension<sup>18,19</sup>. If this property holds



**Figure 3 | Rad55–Rad57 inhibit disruption of Rad51 presynaptic filaments by Srs2.** **a**, RPA–ssDNA complex. **b**, Short (145 nm) Rad51–ssDNA filament. **c**, Long (350 nm) Rad51–ssDNA filaments. **d**, Quantification of electron microscopic analysis. For each reaction condition 300–400 filaments were analysed (2.34  $\mu$ M Rad51, 7  $\mu$ M 600-nucleotide ssDNA,  $\pm$  0.43  $\mu$ M Rad55–Rad57,  $\pm$  0.21  $\mu$ M RPA,  $\pm$  0.4  $\mu$ M Srs2), and the means ( $\phi$ )  $\pm$  1 s.d. and distributions of filament length classes are shown. Scale bars, 100 nm. White arrows indicate RPA–ssDNA complexes and red arrows indicate Rad51 filaments.

true for ssDNA, the formation of a co-filament with Rad51 by Rad55–Rad57 might provide a mechanism to form extended Rad51 filaments. This could also explain the increase in Rad55–Rad57 focus intensity over time after ionizing radiation exposure, and is consistent with the dependence of Rad55–Rad57 foci on Rad51 (ref. 10).

Our biochemical data are consistent with a model (Supplementary Fig. 1) whereby Rad51 presynaptic filament formation is modulated by a balance between the stabilizing function of Rad55–Rad57 and the destabilizing function of Srs2 antirecombinase. This model predicts that a deletion of *SRS2* should suppress the phenotypes caused by defects in Rad55–Rad57. In fact, *srs2Δ* completely suppresses the ionizing radiation sensitivity of *rad57* and *rad55* mutations in quantitative survival assays (Supplementary Fig. 12), consistent with semiquantitative results using *rad57* (ref. 20). However, *srs2Δ* only mildly suppresses the methyl methanesulphonate sensitivity (Supplementary Fig. 13) and recombination defect (Supplementary Fig. 14) of a *rad55* mutation, consistent with previous *rad57* data<sup>20</sup>. The difference in suppression is likely related to a difference in substrates: ionizing radiation-induced DNA damage requires primarily double-strand break repair, whereas methyl



**Figure 4 | Rad55–Rad57 interact with Srs2 and inhibit Srs2 helicase.** **a**, Pull-down with 4 nM (1 pmol) Rad55–Rad57 and 2.7, 8, or 16 nM Srs2. **b**, Pull-down with 4 nM Rad55–Rad57 and 8 or 16 nM Rad51. **c**, Quantification of results in **a** and **b** and additional experiments. **d**, Pull-down with 4 nM Rad55–Rad57 and 8 nM Srs2  $\pm$  40 nM Rad51. GST was used as control. S, supernatant; W, wash; E, eluate. **e**, Helicase assay. **f**, Rad51 (28 nM) with or without Rad55–Rad57 (25 nM) were incubated with 1.5 nM 3'-tailed substrate before addition of 120 nM Srs2 protein. **g**, Product yields at 20 min. Means  $\pm$  1 s.d.,  $n = 3$ , are shown.

methanesulphonate-induced DNA damage and sister chromatid recombination require gap repair (Supplementary Fig. 1). We propose that the Rad51 presynaptic filament is a meta-stable reversible intermediate, whose dynamics in yeast are partially controlled by the balance of the filament-stabilizing activity of Rad55–Rad57 and the filament-destabilizing activity of the Srs2 helicase (Supplementary Fig. 1). This balance is likely to be influenced by the multiple post-translational modifications that have been identified to regulate Rad55–Rad57 (ref. 21) and Srs2 (ref. 22) functions (Supplementary Fig. 1). Together with the local availability of SUMO–PCNA, which specifically recruits Srs2 (refs 23–25), post-translational modifications may determine the balance between recombination and antirecombination in wild-type cells and explain the various degrees of suppression observed in the *srs2 rad55* (*rad57*) double mutants that depend on the type of DNA damage or genetic endpoint (double-strand break versus replication-fork-associated gap in Supplementary Fig. 1).

The human RAD51 paralogues have important roles in tumour suppression and human disease<sup>3,26</sup>. Our studies established an unprecedented mechanism of anti-antirecombination that may serve as a

paradigm for the mechanism of action of the five human RAD51 paralogues. The diversification of the human RAD51 paralogues may reflect the multiplicity of human motor proteins that may disrupt RAD51 presynaptic filaments, including the RecQ-like helicases BLM and RECQL5 as well as FBH1 and FANCI<sup>27–30</sup> or indicate additional functions during recombinational repair.

## METHODS SUMMARY

Purification of yeast Rad51, Rad55–Rad57, RPA and Srs2, the biochemical assays and the electron microscopy analysis are detailed in Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 25 September 2009; accepted 30 August 2011.**

**Published online 23 October 2011.**

- Heyer, W. D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **44**, 113–139 (2010).
- Symington, L. S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.* **66**, 630–670 (2002).
- Thacker, J. The RAD51 gene family, genetic instability and cancer. *Cancer Lett.* **219**, 125–135 (2005).
- Sung, P. Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. *Genes Dev.* **11**, 1111–1121 (1997).
- Lovett, S. T. & Mortimer, R. K. Characterization of null mutants of the RAD55 gene of *Saccharomyces cerevisiae*: effects of temperature, osmotic strength and mating type. *Genetics* **116**, 547–553 (1987).
- Beernink, H. T. H. & Morrical, S. W. RMPs: recombination/replication mediator proteins. *Trends Biochem. Sci.* **24**, 385–389 (1999).
- Morimatsu, K. & Kowalczykowski, S. C. RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol. Cell* **11**, 1337–1347 (2003).
- Liu, J., Doty, T., Gibson, B. & Heyer, W. D. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nature Struct. Mol. Biol.* **17**, 1260–1262 (2010).
- Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678–683 (2010).
- Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699–713 (2004).
- Schiestl, R. H., Prakash, S. & Prakash, L. The SRS2 suppressor of *rad6* mutations of *Saccharomyces cerevisiae* acts by channeling DNA lesions into the RAD52 DNA repair pathway. *Genetics* **124**, 817–831 (1990).
- Aboussekhra, A. et al. RADH, a gene of *Saccharomyces cerevisiae* encoding a putative DNA helicase involved in DNA repair. Characteristics of *radH* mutants and sequence of the gene. *Nucleic Acids Res.* **17**, 7211–7219 (1989).
- Aguilera, A. & Klein, H. L. Genetic control of intrachromosomal recombination in *Saccharomyces cerevisiae*. I. Isolation and genetic characterization of hyper-recombination mutations. *Genetics* **119**, 779–790 (1988).
- Krejci, L. et al. DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423**, 305–309 (2003).
- Veaute, X. et al. The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature* **423**, 309–312 (2003).
- Antony, E. et al. Srs2 disassembles Rad51 filaments by a protein-protein interaction triggering ATP turnover and dissociation of Rad51 from DNA. *Mol. Cell* **35**, 105–115 (2009).
- Dupaigne, P. et al. The Srs2 helicase activity is stimulated by Rad51 filaments on dsDNA: implications for crossover incidence during mitotic recombination. *Mol. Cell* **29**, 243–254 (2008).
- Hilario, J., Amitani, I., Baskin, R. J. & Kowalczykowski, S. C. Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. *Proc. Natl Acad. Sci. USA* **106**, 361–368 (2009).
- Modesti, M. et al. Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. *Structure* **15**, 599–609 (2007).
- Fung, C. W., Mozlin, A. M. & Symington, L. S. Suppression of the double-strand-break-repair defect of the *Saccharomyces cerevisiae* rad57 mutant. *Genetics* **181**, 1195–1206 (2009).
- Herzberg, K. et al. Phosphorylation of Rad55 on serines 2, 8, and 14 is required for efficient homologous recombination in the recovery of stalled replication forks. *Mol. Cell. Biol.* **26**, 8396–8409 (2006).
- Saponaro, M. et al. Cdk1 targets Srs2 to complete synthesis-dependent strand annealing and to promote recombinational repair. *PLoS Genet.* **6**, e1000858 (2010).
- Papouli, E. et al. Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. *Mol. Cell* **19**, 123–133 (2005).
- Pfander, B., Moldovan, G. L., Sacher, M., Hoege, C. & Jentsch, S. SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* **436**, 428–433 (2005).
- Burgess, R. C. et al. Localization of recombination proteins and Srs2 reveals anti-recombinase function *in vivo*. *J. Cell Biol.* **185**, 969–981 (2009).
- Meindl, A. et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nature Genet.* **42**, 410–414 (2010).
- Hu, Y. et al. RECQL5/Recq15 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments. *Genes Dev.* **21**, 3073–3084 (2007).
- Bugreev, D. V., Yu, X., Egelman, E. H. & Mazin, A. V. Novel pro- and anti-recombination activities of the Bloom's syndrome helicase. *Genes Dev.* **21**, 3085–3094 (2007).
- Sommers, J. A. et al. FANCI uses its motor ATPase to destabilize protein-DNA complexes, unwind triplexes, and inhibit RAD51 strand exchange. *J. Biol. Chem.* **284**, 7505–7517 (2009).
- Fugger, K. et al. Human Fbh1 helicase contributes to genome maintenance via pro- and anti-recombinase activities. *J. Cell Biol.* **186**, 655–663 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Alexeeva for the cell culture support. We thank P. Sung, R. Kolodner and L. Symington for plasmids and yeast strains. We are grateful to S. Kowalczykowski, N. Hunter, D. Castaño-Diez, P. Ringler and all members of the Heyer laboratory for discussions and comments on the manuscript. This work was supported by Postdoctoral Fellowship 17FT-0046 from the Tobacco-Related Disease Research Program (J.L.), by the European Community (LSHG-CT-2003-503303) and the Centre National de la Recherche Scientifique, the Commissariat à l'Energie Atomique (X.V., F.F.), by SystemsX.ch (H.S.), and the National Institutes of Health grants U54GM74929 (H.S.), CA92267 and GM58015 (W.-D.H.).

**Author Contributions** J.L. designed, performed and analysed all experiments, except the ionizing radiation survival assay, and helped write the manuscript. L.R. helped with the electron microscopy image collection and data analysis. X.V. purified the Srs2 protein. F.F. performed the ionizing radiation experiment. H.S. advised on the electron microscopy analysis. W.-D.H. conceived the project, designed experiments, coordinated collaborations, contributed to data analysis and wrote the manuscript. All authors discussed results and edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to W.-D.H. ([wdheyer@ucdavis.edu](mailto:wdheyer@ucdavis.edu)).

## METHODS

**Protein purification.** Yeast Rad51, RPA and Srs2 proteins were purified as described<sup>15,31</sup>. The purification of Rad55–Rad57 was adapted from a previously published protocol<sup>32</sup>. Yeast cells overexpressing GST–Rad55–His<sub>6</sub>–Rad57 were grown and harvested as described<sup>32</sup>. Cells were disrupted in Buffer B containing 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, 2.7 mM KCl, 1 M NaCl, 10% (v/v) glycerol, 10 mM β-mercaptoethanol and protease inhibitor cocktail (1 mM PMSF, 2 μM leupeptin, 1 μM pepstatin A and 1 mM benzamide) using glass beads (0.5 mm glass beads; BioSpec Products, Inc.). The cell lysate was centrifuged at 40,000 r.p.m. for 45 min using a Ti50.2 rotor. The supernatant was collected and loaded onto a pre-equilibrated Glutathione Sepharose 4B column (GE Healthcare). After washing with buffer A (20 mM Tris–HCl pH 7.5, 1 mM EDTA, 1 M NaCl, 5 mM β-mercaptoethanol and 10% glycerol), the GST-tagged proteins were eluted with Buffer A containing 20 mM reduced glutathione plus protease inhibitor cocktail. Fractions containing the GST–Rad55–His<sub>6</sub>–Rad57 heterodimer, as determined by 10% SDS–PAGE, were pooled and dialysed against Buffer C (50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8.0, 1 M NaCl and 10% glycerol) containing the protease inhibitor cocktail. Then the pool was loaded onto a pre-equilibrated Ni–NTA agarose column and washed with Buffer C plus protease inhibitor cocktail. The bound complexes were eluted with Buffer C containing 0.5 M NaCl, 0.1 mM PMSF and 250 mM imidazole, and analysed by 10% SDS–PAGE. Fractions containing stoichiometric Rad55–Rad57 heterodimer were pooled, concentrated, dialysed into the storage buffer containing 20 mM Tris–HCl pH 7.5, 0.5 M NaCl, 0.1 mM EDTA, 1 mM DTT and 10% glycerol, and then stored in aliquots at –80 °C. The absence of contaminating enzymatic activities and DNA in protein preparations was verified as described<sup>33</sup>.

**Purification of 600-nucleotide ssDNA.** As published previously<sup>17</sup>, 600-bp dsDNA fragments biotinylated at a one 5' end were generated by PCR from PstI-linearized φX174 DNA using primers WDHY427 5'-TTATCGAAGCGCGCATAAAT-3' and 5' biotinylated WDHY431 5'-GTCCTTCATTTCCATGCGG TG-3'. The biotinylated dsDNA was loaded onto a HiTrap Streptavidin HP column (Amersham Biosciences), and non-biotinylated single-stranded DNA was eluted with 60 mM NaOH.

**Rad51–ssDNA filament assembly assay.** In Fig. 1b, Rad51 (0.67 μM) was incubated with 4 μM ssDNA, in the presence or absence of 0.11 μM Rad55–Rad57, in buffer R containing 20 mM triethanolamine pH 7.5, 4 mM magnesium acetate, 2.5 mM ATP, 25 μg ml<sup>–1</sup> BSA, 1 mM DTT, 90 mM NaCl and 5% glycerol for 10 min. Then 0.25% glutaraldehyde was used to crosslink the protein–DNA complexes for 15 min. The complexes were separated on a 0.5% agarose gel, stained with SYBR Gold, transferred to nitrocellulose membrane, and blotted with rabbit polyclonal anti-Rad51 or anti-Rad55 antibodies.

**Rad51–ssDNA filament salt challenge assay.** In Supplementary Fig. 3, Rad51 (0.267 μM) was incubated with 4 μM ssDNA, in the presence or absence of 0.067 μM Rad55–Rad57, in buffer R containing 20 mM triethanolamine pH 7.5, 4 mM magnesium acetate, 2.5 mM ATP, 25 μg ml<sup>–1</sup> BSA, 1 mM DTT and 5% glycerol for 10 min. Then 5 M stock NaCl solution was added to the reaction to reach a final concentration of 500 mM for a further incubation of 30 min. Glutaraldehyde (0.25%) was used to crosslink the protein–DNA complexes for 15 min. Complexes were separated on a 0.5% agarose gel and stained with SYBR Gold. Proteins were transferred to nitrocellulose membrane and blotted with anti-Rad51 antibodies. All DNA concentrations refer to nucleotides (ssDNA) or base pairs (dsDNA).

**Protein binding to ssDNA immobilized on magnetic beads.** In Fig. 2b, a 5'-biotinylated oligonucleotide was immobilized onto magnetic streptavidin beads as previously described<sup>34</sup>. The oligo sequence is 5'-CCCCCCCCCCCCCAAGATAATTTTTCGACTCATCAGAAATATCCGAAAGTGTTAACTTCTGCGTCATGGAAGCGATAAACTC-3'. In experiments containing Srs2, 10-μl slurry of beads containing 3 μM ssDNA was incubated with 1 μM Rad51 in the presence and absence of 0.1 μM Rad55–Rad57 in buffer containing 20 mM triethanolamine, 5 mM magnesium acetate, 4 mM ATP, 25 μg ml<sup>–1</sup> BSA, 1 mM DTT, 5% glycerol and 25 mM NaCl for 10 min at 22 °C. Then 0.1 or 0.33 μM Srs2 protein was added and further incubated for 10 min. The beads were washed, and bound proteins were eluted and quantified as described<sup>35</sup>. Background protein binding was typically less than 3%.

**Topology-based assay for Rad51 dissociation.** A published protocol was modified slightly for this assay using M13mp18 ssDNA<sup>36</sup>. In Supplementary Fig. 4, 375 nM Rad51, with 0, 80 and 120 nM Rad55–Rad57, were incubated with 9 μM circular M13mp18 ssDNA in 25 μl of buffer containing 20 mM triethanolamine pH 7.5, 4 mM magnesium acetate, 25 μg ml<sup>–1</sup> BSA, 1 mM DTT, and an ATP-regenerating system consisting of 2.5 mM ATP, 20 U ml<sup>–1</sup> creatine kinase, and 20 mM creatine phosphate for 10 min at 30 °C. Then, 100 nM Srs2 and 150 nM RPA were added and incubated for 10 min, before the addition of topologically relaxed pUC19 dsDNA (7 μM in base pairs) and wheat germ DNA topoisomerase I (3 U). After

another 10 min incubation, reactions were stopped by addition of 4 μl stop solution consisting of 1% (w/v) SDS, 75 mM EDTA, 10 mg ml<sup>–1</sup> protease K and further 30 min incubation at 37 °C. DNA species were resolved by electrophoresis on a 1% TBE-agarose gel and visualized using ultraviolet transillumination after ethidium bromide staining. The results were quantified using ImageQuant.

**Protein interaction assays.** GST–Rad55–His<sub>6</sub>–Rad57 (4 nM) or 30 nM GST (GE Healthcare) were incubated with increasing amounts of either Srs2 or Rad51 in buffer P containing 25 mM Tris–HCl (pH 7.5), 10 mM magnesium acetate, 50 mM NaCl, 1 mM DTT, 10% glycerol and 0.05% NP-40 for 1 h at room temperature (Fig. 4a–d). Equilibrated and BSA-treated Glutathione-Sepharose 4B beads were added to the mixture and incubated for 1 h. The beads and supernatant were separated by centrifugation and the beads were washed twice with binding buffer P. The pulled-down protein complexes were eluted by boiling at 95 °C for 3 min in 10 μl SDS–PAGE loading buffer, separated through a 10% SDS–PAGE gel, and the protein bands were visualized through immunoblots and quantified by ImageQuant. In Fig. 4a, b, 1/16th of the supernatant and wash were loaded. In Fig. 4d, 1/7th of the supernatant and wash were loaded. For the competition protein binding assay (Supplementary Figs 7 and 8), the two proteins were incubated for 30 min before the addition of an increasing amount of the third challenging protein, as specified in the diagrams. After another 30-min incubation, equilibrated and BSA-treated Glutathione-Sepharose 4B beads were added to the mixture and incubated for 1 h. Analysis and quantification was performed as described above. The anti-Rad51, –Rad55 and –Rad57 antibodies were generated in rabbits, the anti-Srs2 antibody was purchased from Santa Cruz Biotechnology.

**Helicase assay.** The assay followed a published protocol and the substrates were prepared exactly as described before<sup>15</sup>. In Fig. 4f, 28 nM Rad51 with 0 or 25 nM Rad55–Rad57 were incubated with 1.5 nM oligo substrate with 3' tail in buffer containing 20 mM triethanolamine pH 7.5, 4 mM magnesium acetate, 25 μg ml<sup>–1</sup> BSA, an ATP-regenerating system consisting of 2.5 mM ATP, 20 U ml<sup>–1</sup> creatine kinase and 20 mM creatine phosphate, as well as either 1 mM DTT and 40 mM NaCl (Fig. 4e, g and Supplementary Fig. 10c, d) or 5 mM DTT and 10 mM NaCl (Supplementary Fig. 10a) for 10 min at 30 °C. Then 120 nM Srs2 protein was added to initiate the helicase reaction. After 20 min incubation, the reactions were stopped by adding 4.5 μl stop buffer containing 150 mM EDTA, 2% SDS, 163 nM unlabelled oligo, and 4.3 mg ml<sup>–1</sup> protease K into 9 μl reaction sample. The DNA species were separated through electrophoresis on a 10% TBE–PAGE gel, which was dried and analysed by a Storm phosphorimager. The bands were quantified by densitometry using ImageQuant.

**DNA strand exchange assay.** In Supplementary Fig. 11, Rad51 (3.3 μM) was incubated with 0.3 μM Rad55–Rad57 or the corresponding amount of Rad55–Rad57 storage buffer and 10 μM φX174 ssDNA for 15 min at 30 °C in buffer containing 30 mM Tris-acetate (pH 7.5), 4 mM magnesium acetate, 75 mM NaCl, 1 mM DTT, 2.5 mM ATP, 50 μg ml<sup>–1</sup> BSA, 20 mM phosphocreatine and 80 ng μl<sup>–1</sup> creatine kinase. 0.56 μM RPA and 0, 333, 222, 167, 125 nM of Srs2 were added, and incubated for another 30 min. Then 10 μM (bp) PstI-linearized φX174 dsDNA and 4.8 mM spermidine were added and further incubated for 120 min. Samples were deproteinized and separated by electrophoresis on a 0.8% TBE-agarose gel. Images were recorded using a FluorChem8900 imaging system (Alpha Innotech) after staining with SYBR-Gold (Invitrogen), and quantified with ImageQuant. Percentage of joint molecule (JM) was calculated according to the equation  $JM\% = (JM/1.5)/(JM/1.5 + NC + dsDNA)$ . Percentage of product formation was calculated according to the equation  $product\% = (JM/1.5 + NC)/(JM/1.5 + NC + dsDNA)$ . NC, nicked circle.

**Electron microscopy.** To assemble the protein–DNA filament, 2.34 μM Rad51 protein, in the presence or absence of 0.43 μM Rad55–Rad57, was incubated with 7 μM 600-nt ssDNA (+) strand for 10 min at 30 °C in 20 mM triethanolamine pH 7.5, 4 mM magnesium acetate, 1 mM DTT and 3 mM ATP. RPA (0.21 μM) was added and incubated for another 10 min. Lastly, 0.4 μM Srs2 or buffer control was added and incubated for 10 min. The reaction mixtures were diluted 20-fold in 10 mM Tris–HCl pH 7.5, 50 mM NaCl and 5 mM MgCl<sub>2</sub> without chemical fixation. The samples were adsorbed onto 400 mesh carbon-coated copper grids (Ted Pella), negatively stained with 2% (w/v) uranyl acetate, blotted, and air-dried. Grids were imaged in a JEOL JEM-1230 transmission electron microscope (JEOL). Images of Rad51-filaments were randomly collected from different areas on the grid. 6–10 grids were used for each condition. Images were recorded at a nominal magnification of ×40,000 under minimum dose procedures on a Tietz 2,048 × 2,048 pixel CCD camera (TVIPS, Germany). Immunoaffinity gold labelling of GST–Rad55, as shown in Fig. 1d, was adapted from a published protocol<sup>37</sup>. In brief, Rad51–Rad55–Rad57–ssDNA complexes were assembled as described above and crosslinked with 0.25% glutaraldehyde for 20 min, before deposition on grids. Grids were blocked in 50 μg ml<sup>–1</sup> BSA in TBST for 30 min, and then incubated with goat anti-GST antibody (GE Healthcare) for 30 min. After three 5 min washes with 50 μg ml<sup>–1</sup> BSA in TBST, the grids were incubated in TBST plus a 1:5

dilution of gold particles dressed with rabbit anti-goat antibody (BioAssay Works). After two 5 min washes in 50  $\mu\text{g ml}^{-1}$  BSA in TBST and one 5 min wash in 5 mM magnesium acetate, grids were stained with 2% (w/v) uranyl acetate before imaging.

***Saccharomyces cerevisiae* strains.** Strains used are listed in Supplementary Table 2.

**Recombination assay.** Spontaneous recombination rates between direct repeats were determined following a published fluctuation analysis protocol using the method of the median<sup>38,39</sup>. The direct-repeat recombination substrate has two different *ade2* alleles separated by plasmid sequences and the *URA3* gene<sup>20</sup>. Yeast strains were grown on YPD plates for 2 days at 30 °C for single colonies. For each strain, nine independent single colonies were randomly chosen and the entire colony was used to inoculate 4 ml YPD liquid culture. Liquid cultures were grown for 2–3 days at 30 °C to reach stationary phase. Cells were collected, washed with sterile H<sub>2</sub>O, and suspended into 1 ml sterile H<sub>2</sub>O. 100  $\mu\text{l}$  of appropriate dilutions of each culture were spread on two plates each of SD-ADE-URA. Cells were incubated for 2 days at 30 °C. For each culture, the number of colonies on YPD were counted and totalled to determine the total cell number. The number of colonies on SD-ADE-URA were counted to determine the median number of recombinants. For each strain, recombination rates were measured independently three times and the mean values with standard deviations are shown.

**MMS sensitivity assay.** Yeast strains were grown overnight in liquid YPD to mid-log phase at 30 °C, and then diluted to OD<sub>600 nm</sub> = 1. Serial dilutions of these cell cultures were made with sterile H<sub>2</sub>O and spotted onto YPD plates with or without methyl methanesulphonate. Plates were incubated for 3 days at 30 °C or 5 days at 22 °C before photographing using a FluorChem8900 imaging system (Alpha Innotech).

**Ionizing radiation survival assay.** Exponentially growing cells ( $1 \times 10^7$  to  $2 \times 10^7$  per ml) in YPD medium at 28 °C were collected by centrifugation, washed

in cold saline (0.9% NaCl), sonicated and resuspended in saline at the desired concentration. The cell suspension was  $\gamma$ -irradiated in a <sup>137</sup>Cs irradiator delivering 20 Gy min<sup>-1</sup>. Aliquots of appropriate dilutions were spread on YPD-containing plates pre-warmed at either 23 °C or 34 °C. The plates were incubated at the corresponding temperature for 4 days (34 °C) or 6 days (23 °C) before counting the colonies. Platings were done in duplicate. The experiments were repeated at least three times, and the result of one typical assay is shown.

31. Solinger, J. A., Kiianitsa, K. & Heyer, W.-D. Rad54, a Swi2/Snf2-like recombinational repair protein, disassembles Rad51:dsDNA filaments. *Mol. Cell* **10**, 1175–1188 (2002).
32. Bashkurov, V. I., Herzberg, K., Haghazari, E., Vlasenko, A. S. & Heyer, W. D. DNA-damage induced phosphorylation of Rad55 protein as a sentinel for DNA damage checkpoint activation in *S. cerevisiae*. *Methods Enzymol.* **409**, 166–182 (2006).
33. Zhang, X. P. & Heyer, W. D. Quality control of purified proteins involved in homologous recombination. *Methods Mol. Biol.* **745**, 329–343 (2011).
34. Mazin, A. V., Alexeev, A. A. & Kowalczykowski, S. C. A novel function of Rad54 protein - Stabilization of the Rad51 nucleoprotein filament. *J. Biol. Chem.* **278**, 14029–14036 (2003).
35. Wolner, B., van Komen, S., Sung, P. & Peterson, C. L. Recruitment of the recombinational repair machinery to a DNA double-strand break in yeast. *Mol. Cell* **12**, 221–232 (2003).
36. Schwendener, S. et al. Physical interaction of RECQ5 helicase with RAD51 facilitates its anti-recombinase activity. *J. Biol. Chem.* **285**, 15739–15745 (2010).
37. Van Dyck, E., Hajibagheri, N. M., Stasiak, A. & West, S. C. Visualisation of human RAD52 protein and its complexes with hRad51 and DNA. *J. Mol. Biol.* **284**, 1027–1038 (1998).
38. Lea, D. E. & Coulson, C. A. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **49**, 264–285 (1949).
39. Spell, R. M. & Jinks-Robertson, S. Determination of mitotic recombination rates by fluctuation analysis in *Saccharomyces cerevisiae*. *Methods Mol. Biol.* **262**, 3–12 (2004).

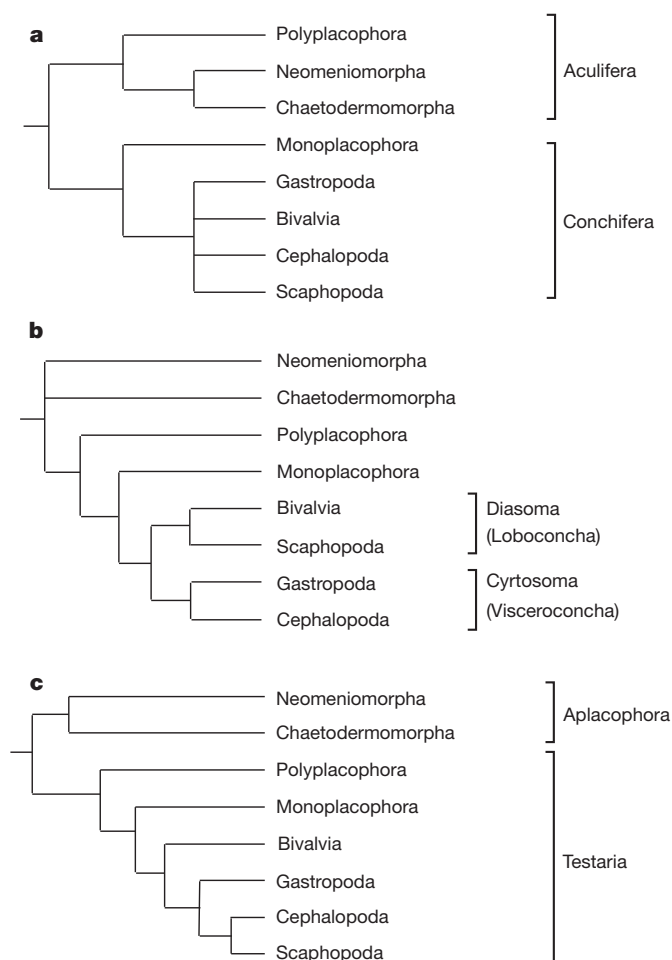
# Resolving the evolutionary relationships of molluscs with phylogenomic tools

Stephen A. Smith<sup>1,2</sup>, Nerida G. Wilson<sup>3,4</sup>, Freya E. Goetz<sup>1</sup>, Caitlin Feehery<sup>1,4</sup>, Sónia C. S. Andrade<sup>5</sup>, Greg W. Rouse<sup>4</sup>, Gonzalo Giribet<sup>5</sup> & Casey W. Dunn<sup>1</sup>

Molluscs (snails, octopuses, clams and their relatives) have a great disparity of body plans and, among the animals, only arthropods surpass them in species number. This diversity has made Mollusca one of the best-studied groups of animals, yet their evolutionary relationships remain poorly resolved<sup>1</sup>. Open questions have important implications for the origin of Mollusca and for morphological evolution within the group. These questions include whether the shell-less, vermiform aplacophoran molluscs diverged before the origin of the shelled molluscs (Conchifera)<sup>2–4</sup> or lost their shells secondarily. Monoplacophorans were not included in molecular studies until recently<sup>5,6</sup>, when it was proposed that they constitute a clade named Serialia together with Polyplacophora (chitons), reflecting the serial repetition of body organs in both groups<sup>5</sup>. Attempts to understand the early evolution of molluscs become even more complex when considering the large diversity of Cambrian fossils. These can have multiple dorsal shell plates and sclerites<sup>7–10</sup> or can be shell-less but with a typical molluscan radula and serially repeated gills<sup>11</sup>. To better resolve the relationships among molluscs, we generated transcriptome data for 15 species that, in combination with existing data, represent for the first time all major molluscan groups. We analysed multiple data sets containing up to 216,402 sites and 1,185 gene regions using multiple models and methods. Our results support the clade Aculifera, containing the three molluscan groups with spicules but without true shells, and they support the monophyly of Conchifera. Monoplacophora is not the sister group to other Conchifera but to Cephalopoda. Strong support is found for a clade that comprises Scaphopoda (tusk shells), Gastropoda and Bivalvia, with most analyses placing Scaphopoda and Gastropoda as sister groups. This well-resolved tree will constitute a framework for further studies of mollusc evolution, development and anatomy.

Since the first animal phylogenies based on molecular data, many researchers have struggled to resolve mollusc phylogenies even as taxon sampling improved<sup>5,6,12</sup> (see Fig. 1 for some hypotheses that have been proposed). Little support, if any, was found for the monophyly of Mollusca or most of its larger subclades. Better results were achieved for some internal relationships of these groups, including Polyplacophora, Bivalvia, Cephalopoda, Scaphopoda and Gastropoda, although often with difficulties recovering monophyly of the two largest clades, the gastropods and bivalves<sup>5,13,14</sup>. Unfortunately, fundamental questions in mollusc evolution remain largely unanswered by the molecular and morphological data. These questions include whether the aplacophoran molluscs are monophyletic<sup>2</sup> or paraphyletic<sup>3,4</sup>. There has also been conflicting evidence for the placement of Polyplacophora, which has been placed with the aplacophorans (forming the clade Aculifera), as the sister group to the shelled molluscs (forming the clade Testaria) or as the sister group to Monoplacophora (forming the clade Serialia). In addition, many hypotheses have been proposed for the interrelationships of the conchiferan groups. The extensive fossil record of

Mollusca (which dates back to the Cambrian), combined with the numerous Palaeozoic forms that are considered stem-group molluscs and the lack of resolution in targeted-gene approaches to molluscan phylogenetics, pointed towards a possible rapid radiation with little phylogenetic signal left in the genomes of molluscs. However, the same has been argued for the radiation of Metazoa in the Cambrian or earlier<sup>15</sup>, but large increases in gene representation using phylogenomic analyses have clearly ameliorated this problem and identified relationships that seemed impossible to resolve with target-gene



**Figure 1 | Selected hypotheses of extant molluscan relationships and relevant taxa.** Phylograms based on the hypotheses of Scheltema<sup>3</sup> (a), Salvini-Plawen and Steiner<sup>2</sup> (b) and Waller<sup>28</sup> (c). Most controversy centres on the monophyly of Aplacophora, the relationships within Conchifera and the placement of Polyplacophora (for example, in Aculifera versus in Testaria).

<sup>1</sup>Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island 02912, USA. <sup>2</sup>Heidelberg Institute for Theoretical Studies, Heidelberg D-69118, Germany. <sup>3</sup>The Australian Museum, Sydney, New South Wales 2010, Australia. <sup>4</sup>Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA. <sup>5</sup>Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

approaches<sup>16,17</sup>. We applied the same principles to Mollusca, one of the most challenging problems to solve in animal phylogenetics.

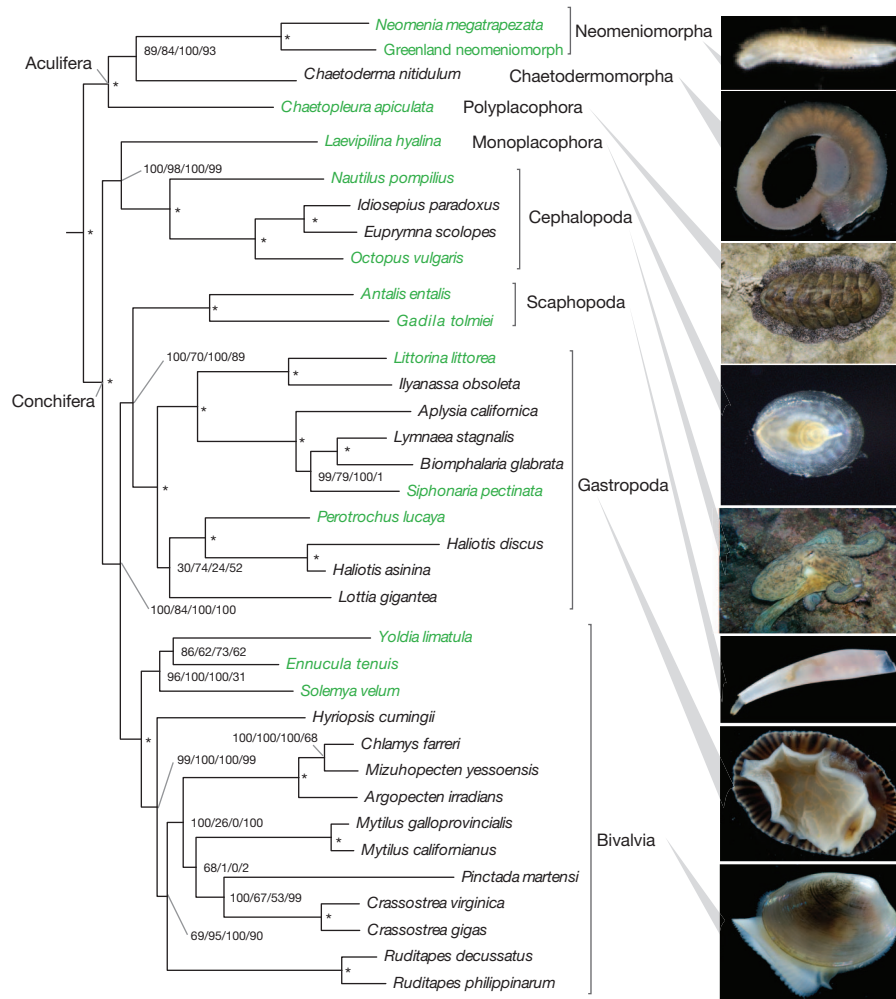
Only phylogenomic analyses have been able to recover molluscan monophyly with high support<sup>17</sup>; however, few molluscs were included in earlier analyses, and not all of the major subclades were represented. Therefore, little could be concluded about the interrelationships of the major molluscan groups. Morphology-based cladistic analyses have often relied on 'idealized' composite ground patterns to represent entire clades<sup>2,4</sup>, a practice that has now been largely replaced with the use of exemplar species<sup>18</sup> and more detailed character descriptions. But an analysis of molluscan morphological features coding real exemplars has yet to be published, and the exemplar approach is much more amenable to molecular data.

Analyses of our broadly sampled, new phylogenomic data (see Methods, Supplementary Table 1 and Supplementary Fig. 1) result in a well-resolved and highly supported phylogeny of Mollusca (Fig. 2), in contrast to all previous molecular attempts<sup>5,6,12</sup>. These results are consistent across analytical methods, phylogenetic inference programs, matrices that vary in occupancy and the number of genes considered (Fig. 2 and Supplementary Figs 2–6 and 9), and the inclusion of different outgroup taxa (Supplementary Fig. 7).

Our results (Fig. 2) show a sister group relationship between the aculiferan molluscs and the conchiferan groups. Aculifera<sup>3,19</sup> includes

Polyplacophora as the sister group to the two aplacophoran groups (Neomeniomorpha and Chaetodermomorpha). This topology lends support to the idea that the vermiform Aplacophora are not plesiomorphic but are derived from plated Palaeozoic molluscs such as *Acaenoplax*<sup>10</sup>. The aculiferans are characterized by spicules and dorsal shell plates. Chitons have eight dorsal shell plates, but their larva has an anlagen with seven rows of dorsal papillae, as observed in the serially arranged spiculoblasts of a chaetodermomorph larva<sup>20</sup>, a character that may constitute a synapomorphy of the clade.

Conchifera is supported as a clade, suggesting that true shells may have originated only once, perhaps by the concentration of a diffuse shell gland into a single zone of the mantle (two zones in bivalves), at least as defined by the role of *engrailed* during organogenesis<sup>21</sup>. The support here for Conchifera rejects the recent Serialia hypothesis<sup>5,6</sup>. Comparing the site likelihoods in analyses in which Serialia is constrained with those in which it is not constrained reveals that there are many more characters that are incongruent with Serialia than support Serialia (Supplementary Fig. 8). Monoplacophora is not, however, the sister group of all other Conchifera, as has been suggested by most authors, and is instead the sister group to Cephalopoda, as has been proposed based on some palaeontological data<sup>22</sup>. Many palaeontologists have accepted the monoplacophoran 'ancestry' of Cephalopoda<sup>23,24</sup>, although this relationship has been rejected by neontologists, who



**Figure 2 | Phylogram of the RAxML maximum likelihood analysis of the big matrix (216,402 amino acids) under the WAG+ $\Gamma$  model.** Support values for the topology obtained from four analyses are listed as percentages in the order A/B/C/D. A is the bootstrap support from RAxML analysis under the WAG model for the big matrix. B is the bootstrap from RAxML analysis under the

WAG model for the small matrix. C is the posterior probability from MrBayes under the WAG model for the small matrix. D is the posterior probability from PhyloBayes under the CAT model for the small matrix. Asterisks indicate 100/100/100/100 support. Taxa with new data are shown in green. Scale bar, 0.08 expected changes per site.

consider that cephalopods and gastropods share important morphological features such as the presence of cephalic eyes, the isolation of the head from the visceral mass, the terminal position of the mantle cavity and the occurrence of muscle antagonistic systems<sup>2,23</sup>. The presence of multiseptate shells in fossil Hypseloconidae monoplacophorans, a character that is found in *Nautilus* and fossil cephalopods, has been interpreted as supporting this relationship between Cephalopoda and Monoplacophora<sup>24</sup>. The presence of two pairs of gills, kidneys and atria in the chambered *Nautilus* has been interpreted as an indication that secondary simplification took place during the early evolution of cephalopods from an ancestor with serially repeated structures<sup>23</sup>. This interpretation and the present trees suggest that the most recent common ancestor of Cephalopoda and Monoplacophora had some serially repeated structures.

The internal resolution of Cephalopoda is in agreement with all of the current hypotheses, with the chambered *Nautilus* forming the sister group of Coleoidea, and also identifies the monophyly of Decapoda<sup>25,26</sup>. Scaphopoda, Gastropoda and Bivalvia form a clade with thick multilayered shells, but this clade has received little attention in the literature<sup>23</sup>. Most morphological hypotheses place scaphopods as the sister group to bivalves in a clade named Diasoma<sup>2,27</sup> and, recently, molecular and developmental data have favoured a cephalopod–scaphopod relationship. Although there is strong support for the placement of Scaphopoda as the sister group to Gastropoda in maximum likelihood analyses of the big matrix (Fig. 2 and Supplementary Figs 2 and 4), maximum likelihood analyses of the small matrix recover this same relationship but with less support (Fig. 2 and Supplementary Figs 3 and 5). Bayesian analyses using the site-heterogeneous CAT model of protein evolution also place Scaphopoda as the sister group to Gastropoda, with a posterior probability of 89% (Supplementary Fig. 9).

Within Bivalvia, maximum likelihood analyses and Bayesian analyses under the Whelan and Goldman (WAG) model support the monophyly of Protobranchia, which includes bivalves with plesiomorphic ctenidia—gills comparable to those of many other molluscs. This contradicts some earlier bivalve phylogenies, based on fewer data, that proposed paraphyly of protobranchs<sup>14</sup> but supports the traditional morphological views<sup>2,28</sup>. Bayesian analyses with the CAT model are consistent with Protobranchia but do not provide strong support for it. The hypertrophied bivalve gill, which is responsible for filter feeding, had a single origin, and organisms with this type of gill constitute the well-recognized clade Autolamellibranchiata. Palaeoheterodonta (the group that includes freshwater pearl mussels) is the sister group to all other autolamellibranchiates, which can be divided into heterodonts and pteriomorphians. This hypothesis is similar to that proposed by some palaeontologists, although additional taxa, especially *Neotrigonia*, *Anomalodesmata* and *Archiheterodonta*, must be included before concluding more about the internal autolamellibranchiate relationships.

Likewise, the internal relationships of Gastropoda, although still limited in their taxonomic representation (the group includes nearly 100,000 living species), support some of the major divisions that are currently accepted<sup>29</sup>. The patellogastropod *Lottia* is either the sister group to Vetigastropoda (as in Thiele's Archaeogastropoda hypothesis) or the sister group to all other gastropods<sup>29</sup>, depending on the data set that is analysed. The former alternative has been recovered in recent molecular analyses of gastropods<sup>30</sup>. The two representatives of Caenogastropoda form a sister clade to the representatives of Heterobranchia, including opisthobranchs and pulmonates, as suggested in all of the modern analyses of gastropod relationships<sup>13,29</sup>.

For the first time, our data and analyses resolve the broad-scale relationships within Mollusca with strong support. This allows us to gain an understanding not only of the relationships of modern molluscs but also of the numerous Palaeozoic forms of molluscs. It also allows us to investigate several key characters that define the group. Molluscs are related to other animals with spiral development and a trochophore larva and have now been shown to share a close ancestor with annelids

and brachiopods<sup>16</sup>, both of which use chaetoblasts to produce chaetae. Spicules and chaetae may share a similar developmental mechanism<sup>17</sup>. Likewise, the appearance of dorsal plates or shells in addition to sclerites is now well documented in halwaxiids, *Acaenoplax* and Polyplacophora. These features are generated by multiple rows of secretory papillae in chiton and aplacophoran larvae. They may be plesiomorphic among molluscs, especially if halwaxiids are interpreted as stem-group molluscs, but they could also be apomorphic for Aculifera. The condensation of such papillae into a single shell gland<sup>21</sup> could be responsible for the origin of the conchiferan shell, arguably the single event that led to the extraordinary success of molluscs, first in the Cambrian oceans and later in many limnic and terrestrial environments. In addition to the presence of shell glands that can deposit calcium carbonate, the primitive mollusc may have had a rasping radula and serially repeated ctenidia along the mantle cavity, because both characters appear in the two lineages of extant molluscs, Aculifera and Conchifera, as well as in several extinct Palaeozoic stem molluscs. Like the arthropods, with their hardened exoskeletons, molluscs are true conquerors of our land and waters.

## METHODS SUMMARY

New transcriptome data were collected for 14 mollusc species that had been selected to optimize taxonomic representation (Supplementary Table 1). Collecting efforts included an oceanographic campaign to collect members of the key taxon Monoplacophora. Using several protocols, messenger RNA was extracted, and cDNA samples were sequenced on a 454 Genome Sequencer FLX Titanium (Roche) or a Genome Analyzer IIx (Illumina). After assembly and translation, the sequences from all taxa were compared to each other with BLASTP. These pairwise comparisons were used to cluster genes into homologues using the algorithm MCL. The phylogenetic analyses divided sets of homologues into orthologues, which were aligned, trimmed and concatenated into two supermatrices that differed in the number of genes and the average fraction of genes available for each species. The 'small' matrix consists of 301 genes that are present in at least 20 taxa. This matrix has 50% gene occupancy (that is, sequence data were available for an average of 50% of the genes across the taxa), 27% character occupancy (that is, 27% of the matrix consists of unambiguous amino acid data, with the remainder being missing data or alignment gaps) and is 50,930 sites in length. The 'big' matrix consists of 1,185 genes that are present in at least 15 taxa. This matrix has 40% gene occupancy, 21% character occupancy and is 216,402 sites long. Both matrices contain data for all of the 46 species that were included in the study. The matrices were analysed with the programs RAXML, MrBayes and PhyloBayes to infer relationships.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 7 April; accepted 31 August 2011.**

**Published online 26 October 2011.**

1. Ponder, W. F. & Lindberg, D. R. (eds) *Phylogeny and Evolution of the Mollusca* (Univ. California Press, 2008).
2. Salvini-Plawen, L. V. & Steiner, G. in *Origin and Evolutionary Radiation of the Mollusca* (ed. Taylor, J. D.) 29–51 (Oxford Univ. Press, 1996).
3. Scheltema, A. H. in *Origin and Evolutionary Radiation of the Mollusca* (ed. Taylor, J. D.) 53–58 (Oxford Univ. Press, 1996).
4. Haszprunar, G. Is the Aplacophora monophyletic? A cladistic point of view. *Am. Malacol. Bull.* **15**, 115–130 (2000).
5. Giribet, G. *et al.* Evidence for a clade composed of molluscs with serially repeated structures: monoplacophorans are related to chitons. *Proc. Natl Acad. Sci. USA* **103**, 7723–7728 (2006).
6. Wilson, N. G., Rouse, G. W. & Giribet, G. Assessing the molluscan hypothesis Serialia (Monoplacophora + Polyplacophora) using novel molecular data. *Mol. Phylogenet. Evol.* **54**, 187–193 (2010).
7. Vinther, J. & Nielsen, C. The Early Cambrian *Halkieria* is a mollusc. *Zool. Scr.* **34**, 81–89 (2005).
8. Scheltema, A. H., Kerth, K. & Kuzirian, A. M. Original molluscan radula: comparisons among Aplacophora, Polyplacophora, Gastropoda, and the Cambrian fossil *Wiwaxia corrugata*. *J. Morphol.* **257**, 219–245 (2003).
9. Morris, S. C. & Caron, J. B. Halwaxiids and the early evolution of the lophotrochozoans. *Science* **315**, 1255–1258 (2007).
10. Sutton, M. D., Briggs, D. E. G., Siveter, D. J. & Siveter, D. J. Computer reconstruction and analysis of the vermiform mollusc *Acaenoplax hayae* from the Herefordshire Lagerstätte (Silurian, England), and implications for molluscan phylogeny. *Palaeontology* **47**, 293–318 (2004).
11. Caron, J.-B., Scheltema, A., Schander, C. & Rudkin, D. A soft-bodied mollusc with radula from the Middle Cambrian Burgess Shale. *Nature* **442**, 159–163 (2006).

12. Passamanek, Y. J., Schander, C. & Halanych, K. M. Investigation of molluscan phylogeny using large-subunit and small-subunit nuclear rRNA sequences. *Mol. Phylogenet. Evol.* **32**, 25–38 (2004).
13. Aktipis, S. W., Giribet, G., Lindberg, D. R. & Ponder, W. F. in *Phylogeny and Evolution of the Mollusca* (eds Ponder, W. F. & Lindberg, D. R.) 201–237 (Univ. California Press, 2008).
14. Giribet, G. & Distel, D. L. in *Molecular Systematics and Phylogeography of Mollusks* (eds Lydeard, C. & Lindberg, D. R.) 45–90 (Smithsonian Books, 2003).
15. Rokas, A., Krüger, D. & Carroll, S. B. Animal evolution and the molecular signature of radiations compressed in time. *Science* **310**, 1933–1938 (2005).
16. Hejnol, A. *et al.* Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. R. Soc. B* **276**, 4261–4270 (2009).
17. Dunn, C. W. *et al.* Broad taxon sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749 (2008).
18. Prendini, L. Species or supraspecific taxa as terminals in cladistic analysis? Groundplans versus exemplars revisited. *Syst. Biol.* **50**, 290–300 (2001).
19. Vendrasco, M. J., Wood, T. E. & Runnegar, B. N. Articulated Palaeozoic fossil with 17 plates greatly expands disparity of early chitons. *Nature* **429**, 288–291 (2004).
20. Nielsen, C., Haszprunar, G., Ruthensteiner, B. & Wanninger, A. Early development of the aplousophoran mollusc *Chaetoderma*. *Acta Zool.* **88**, 231–247 (2007).
21. Wanninger, A., Koop, D., Moshel-Lynch, S. & Degnan, B. M. in *Phylogeny and Evolution of the Mollusca* (eds Ponder, W. F. & Lindberg, D. R.) 427–445 (Univ. California Press, 2008).
22. Yochelson, E. L. An alternative approach to the interpretation of the phylogeny of ancient mollusks. *Malacologia* **17**, 165–191 (1978).
23. Runnegar, B. in *Origin and Evolutionary Radiation of the Mollusca* (ed. Taylor, J. D.) 77–87 (Oxford Univ. Press, 1996).
24. Yochelson, E. L., Flower, R. H. & Webers, G. F. The bearing of the new Late Cambrian monoplacophoran genus *Knightoconus* upon the origin of Cephalopoda. *Lethaia* **6**, 275–309 (1973).
25. Lindgren, A. R., Giribet, G. & Nishiguchi, M. K. A combined approach to the phylogeny of Cephalopoda (Mollusca). *Cladistics* **20**, 454–486 (2004).
26. Strugnell, J. & Nishiguchi, M. K. Molecular phylogeny of coleoid cephalopods (Mollusca: Cephalopoda) inferred from three mitochondrial and six nuclear loci: a comparison of alignment, implied alignment and analysis methods. *J. Molluscan Stud.* **73**, 399–410 (2007).
27. Runnegar, B. & Pojeta, J. Jr. Molluscan phylogeny: the paleontological viewpoint. *Science* **186**, 311–317 (1974).
28. Waller, T. R. in *Bivalves: An Eon of Evolution* (eds Johnston, P. A. & Haggart, J. W.) 1–45 (Univ. Calgary Press, 1998).
29. Ponder, W. F. & Lindberg, D. R. Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. *Zool. J. Linn. Soc.* **119**, 83–265 (1997).
30. Aktipis, S. W. & Giribet, G. A phylogeny of Vetigastropoda and other ‘archaeogastropods’: re-organizing old gastropod clades. *Invertebr. Biol.* **129**, 220–240 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This research was supported by the US National Science Foundation through the Systematics Program (awards 0844596, 0844881 and 0844652), the ATOL Program (EF-0531757), EPSCoR (Infrastructure to Advance Life Sciences in the Ocean State, 1004057) and the iPlant Collaborative (0735191). Support was also provided by the Scripps Institution of Oceanography, the University of California Ship Funds and the Museum of Comparative Zoology. Collecting in Greenland was supported by the Carlsberg Foundation. A. Riesgo and J. Harasewych provided tissue samples of *Octopus* and *Perotrochus*, respectively. E. Röttinger assisted with *Nautilus*. C. Palacín allowed us to use an *Octopus vulgaris* photograph. At Brown University, Illumina sequencing was enabled by the Genomics Core Facility, and computational analyses were facilitated by L. Dong and the Center for Computing and Visualization. At Harvard University, Illumina sequencing was enabled by the Bauer Core in the Faculty of Arts and Sciences (FAS) Center for Systems Biology, and analyses were supported by the staff of the Research Computing cluster Odyssey facility in the FAS.

**Author Contributions** C.W.D., G.G. and N.G.W. conceived of and oversaw the study. S.A.S. and C.W.D. designed and implemented the data analyses. N.G.W., G.G. and G.W.R. collected the specimens. F.E.G., S.C.S.A. and C.F. prepared the specimens for sequencing. S.A.S., C.W.D., G.G., G.W.R. and N.G.W. wrote the manuscript. All authors read and provided input into the manuscript and approved the final version.

**Author Information** Illumina and 454 reads have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive under accession number SRA044948. Sanger reads for *Laevipilina hyalina* have been deposited in the NCBI Trace Archive under the sequencing centre name BUDL with TI range 2317135955–2317139410. The assembled data, matrices and trees have been deposited in Dryad (<http://dx.doi.org/10.5061/dryad.24cb8>). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to N.G.W. ([nerida.wilson@austmus.gov.au](mailto:nerida.wilson@austmus.gov.au)), G.G. ([ggiribet@oeb.harvard.edu](mailto:ggiribet@oeb.harvard.edu)) or C.W.D. ([casey\\_dunn@brown.edu](mailto:casey_dunn@brown.edu)).

## METHODS

**Taxon sampling and RNA isolation.** The taxa were selected to optimize taxonomic representation within Mollusca. Collecting efforts included an oceanographic campaign to collect members of the key taxon Monoplacophora<sup>31</sup>. New transcriptome data were collected for one outgroup taxon, *Lingula anatina*, and for 14 other taxa that were broadly sampled across Mollusca (Supplementary Table 1). All tissues were collected fresh and were prepared immediately or preserved for subsequent RNA work. Stored tissue was frozen (at  $-80^{\circ}\text{C}$ ) or added to RNAlater (and frozen at  $-80^{\circ}\text{C}$  or  $-20^{\circ}\text{C}$ ). Total RNA was isolated with TRI Reagent (Invitrogen) and further cleaned up with an RNeasy kit (QIAGEN), including a DNase I digestion step.

**Sequencing.** Samples were sequenced on a 454 Genome Sequencer FLX Titanium (Roche) or a Genome Analyzer IIx (GA IIx, Illumina). The sample preparation protocol and sequencing technology used for each sample is listed in Supplementary Table 1.

All 454 samples were sequenced by 454 Life Sciences on one-eighth of a Titanium flow cell. For five of the 454 samples, RNA was sent to the sequencing facility for library preparation and sequencing according to the standard 454 cDNA protocols (these samples are marked Roche in the Library Protocol column of Supplementary Table 1). *Nautilus pompilius* mRNA was enriched by one round of binding to Dynabeads (Invitrogen); for the other specimens, total RNA was sent to the sequencing centre, where mRNA enrichment was performed. For four of the 454 samples, full-length cDNA was prepared according to a template-switching protocol<sup>32</sup> (these samples are marked TS in the Library Protocol column of Supplementary Table 1). Adaptors were modified to include restriction sites and were removed by cleavage before sequencing. An Mmel site was incorporated into the 3' adaptor (5'-ATT CTA GAG CGC ACC TTG GCC TCC GAC TTT TCT TTT CTT TTT TTT TCT TTT TTT TTT VN-3', where V and N are ambiguous nucleotides), and a SfiI site (5'-AAG CAG TGG TAT CAA CGC AGA GTG GCC ACG AAG GCC GGG-3') or an AsiSI site (5'-AAG CAG TGG TAT CAA CGC AGA GTG CGA TCG CGG G-3') was included in the 5' adaptor. Titanium sequencing reagents were used for all samples. Additional expressed sequence tags for *L. hyalina* were sequenced with Sanger technology according to previously described methods<sup>17</sup>.

Most Illumina samples were prepared with the NEBNext mRNA Sample Prep kit (New England BioLabs), with size selection for 400 base pair (bp) products. These samples were sequenced (paired-end, 104 bp), with one per lane on an Illumina GA IIx at the Genomics Core Facility at Brown University. One sample (marked Fragmentase in the Library Protocol column of Supplementary Table 1) was prepared with a modified NEBNext mRNA protocol, in which the full-length cDNA was fragmented with NEBNext dsDNA Fragmentase (New England BioLabs) instead of the mRNA being fragmented. This sample was sequenced (paired-end, 150 bp) in a single lane on an Illumina GA IIx at the FAS Center for Systems Biology at Harvard University.

**Assembly.** Publicly available data from the NCBI dbEST database were processed with a version of the PartiGene pipeline (version 3.0.5)<sup>33</sup> that had been modified to run without user intervention. Trace Archive data were processed as described previously<sup>16</sup>.

Roche 454 data were assembled with the Newbler GS *De novo* Assembler (version 2.3, Roche) with the flags '-cdna -nrm -nosplit'. In cases in which multiple splice variants (isotigs in Newbler terminology) were produced for a gene (an isogroup in Newbler terminology), a single exemplar splice variant was selected. The selected isotig was the one with the highest geometric mean of reads spanning each splice site between contigs. This roughly corresponds to the most abundant splice variant for the gene. Singletons that were not assembled by Newbler were assembled with CAP3 (version 10/15/07, with the options '-z 1' and '-y 100'). The sequences that were assembled by Newbler, the sequences that were assembled by CAP3 and all singletons that were not assembled by either were used in subsequent analyses.

Illumina data were assembled with Velvet<sup>34</sup> (version 1.0.12) and Oases (version 0.1.15). Insert lengths for Oases were estimated with a 2100 Bioanalyzer (Agilent). Reads that did not have an average quality score of at least 35 were removed. We examined the assemblies over a range of  $k$  values (21–61, in increments of 10). We selected a  $k$  value of 61 for all samples, except for *Octopus* (for which we used 31). As for the 454 assemblies, we selected a single splice variant (transcript in Oases terminology) for each gene (locus in Oases terminology). To accomplish this, we developed a procedure whereby we chose transcripts that were at least 150 nucleotides, had a length of at least 85% of the longest transcript for the gene and had the highest read coverage. We ignored loci that had more than 50 transcripts, as these often appeared to be the result of misassembly.

Assembled data were compared to NCBI's nr protein database with BLASTX, with an  $e$  cutoff of 0.00001. Large data sets were compared to a reduced nr database by masking nr sequences from taxa that do not belong to the clade designated by NCBI Taxon ID 33154 (Fungi/Metazoa group). Nucleotide sequences were

translated with a version of the prot4EST (version 2.3)<sup>35</sup> pipeline that had been modified to run without user intervention, using these BLASTX results.

**Orthology assignment.** The orthology assessment for data set assemblies followed one described previously<sup>16</sup>. All-by-all comparisons were conducted with BLASTP as in ref. 17. Clustering analyses were conducted on these results by using MCL<sup>36</sup>. At the suggestion of recent analyses<sup>37</sup>, we excluded edges with  $-\log_{10}$  BLASTP  $e$  values lower than 20, to reduce spurious cluster connections. We examined cluster composition with inflation parameters between 1.1 and 6 and found that the final cluster composition was not particularly sensitive to different inflation values in this range. We selected an inflation value of 2.1. Clusters with at least four taxa and at least one ingroup taxon were aligned by using MAFFT<sup>38</sup> and trimmed with Gblocks<sup>39</sup>, and maximum likelihood analyses were conducted with RAXML<sup>40</sup>. The assessment of these phylogenies was conducted as in ref. 16. Monophyly masking was conducted to reduce the number of monophyletic sequences from the same taxon to one sequence. The resultant phylogenies were then analysed by an iterative paralogy pruning procedure, by which maximally inclusive subtrees with no more than one sequence per taxon were pruned and retained. FASTA-formatted files were generated from subtrees that were produced by the paralogy pruning procedure. These files were then aligned with MAFFT, trimmed with Gblocks, filtered (alignments with fewer than 150 sites were excluded) and concatenated into the final matrices.

**Phylogenetic analyses.** We constructed two phylogenetic matrices from the translated sequences. The 'small' matrix consists of 301 genes that are present in at least 20 taxa. It has 50% gene occupancy and is 50,930 sites long. The 'big matrix' consists of 1,185 genes that are present in at least 15 taxa. It has 40% occupancy and 216,402 sites. Both matrices contain data for all of the 46 species included in the study.

Maximum likelihood analyses were performed for both matrices by using RAXML (version 7.2.6)<sup>40</sup> with both the Le and Gascuel (LG)<sup>41</sup> and WAG<sup>42</sup> models with each gene region partitioned. Likelihood analyses consisted of first conducting a bootstrap analysis with 200 replicates, which was followed by a thorough maximum likelihood search.

Bayesian analyses of the small matrix were conducted with MrBayes (version 3.1.2)<sup>43</sup> and PhyloBayes (version 3.3b)<sup>44,45</sup>. The big matrix was too large to analyse with these tools. With MrBayes, we conducted two searches each with two runs (four runs and 16 chains total). We allowed MrBayes to estimate the fixed rate model of evolution. Each chain was run for 1,000,000 generations, and convergence was determined with time-series plots and an estimated sample size of tree likelihoods of at least 100. Samples recorded before burn-in were removed, and post-burn-in samples of the runs were combined. We summarized the posterior probabilities of the clades with majority rule consensus trees.

We conducted analyses of the reduced-outgroup small matrix with PhyloBayes (version 3.3b) using the CAT model of evolution<sup>45</sup>. PhyloBayes misidentified the data type of our matrix as DNA, resulting in model misspecification and lack of convergence. We conducted the analyses presented here with a modified version that was forced to read all matrices as protein sequences. Five PhyloBayes runs under the fully parameterized CAT model each converged at around 1,500 cycles (at least 86,000 generations) based on time-series plots of the likelihood scores and number of partitions. The runs were allowed to run for 5,000 cycles for two runs and 2,500 cycles for three runs. The runs estimated 140 ( $\pm 10$ ) categories for the model. We removed pre-burn-in samples and constructed a majority rule consensus tree using all five runs (Supplementary Fig. 9).

- Wilson, N. G. *et al.* Field collection of *Laevipilina hyalina* McLean, 1979 from southern California, the most accessible living monoplacophoran. *J. Molluscan Stud.* **75**, 195–197 (2009).
- Ewen-Campen, B. *et al.* The maternal and early embryonic transcriptome of the milkweed bug *Oncopeltus fasciatus*. *BMC Genomics* **12**, 61 (2011).
- Parkinson, J. *et al.* PartiGene: constructing partial genomes. *Bioinformatics* **20**, 1398–1404 (2004).
- Zerbino, D. R. & Birney, E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
- Wasmuth, J. D. & Blaxter, M. L. prot4EST: translating expressed sequence tags from neglected genomes. *BMC Bioinformatics* **5**, 187 (2004).
- van Dongen, S. A *Cluster Algorithm for Graphs* Technical Report No. INS-R0010 (National Research Institute for Mathematics and Computer Science in the Netherlands, Amsterdam, 2000).
- Apeltsin, L., Morris, J. H., Babbitt, P. C. & Ferrin, T. E. Improving the quality of protein similarity network clustering algorithms using the network edge weight distribution. *Bioinformatics* **27**, 326–333 (2011).
- Kato, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinformatics* **9**, 286–298 (2008).
- Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
- Stamatakis, A. P. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
- Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).

42. Whelan, S. & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699 (2001).
43. Huelsenbeck, J., Larget, B., van der Mark, P., Ronquist, F. & Simon, D. *MrBayes: Bayesian Analysis of Phylogeny* (<http://mrbayes.csit.fsu.edu/index.php>) (2005).
44. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286–2288 (2009).
45. Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109 (2004).

# Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants

Daniel J. Gibbs<sup>1\*</sup>, Seung Cho Lee<sup>2\*</sup>, Nurulhikma Md Isa<sup>1</sup>, Silvia Gramuglia<sup>1</sup>, Takeshi Fukao<sup>2</sup>, George W. Bassel<sup>1</sup>, Cristina Sousa Correia<sup>1</sup>, Françoise Corbineau<sup>3</sup>, Frederica L. Theodoulou<sup>4</sup>, Julia Bailey-Serres<sup>2</sup> & Michael J. Holdsworth<sup>1</sup>

Plants and animals are obligate aerobes, requiring oxygen for mitochondrial respiration and energy production. In plants, an unanticipated decline in oxygen availability (hypoxia), as caused by roots becoming waterlogged or foliage submergence, triggers changes in gene transcription and messenger RNA translation that promote anaerobic metabolism and thus sustain substrate-level ATP production<sup>1</sup>. In contrast to animals<sup>2</sup>, oxygen sensing has not been ascribed to a mechanism of gene regulation in response to oxygen deprivation in plants. Here we show that the N-end rule pathway of targeted proteolysis acts as a homeostatic sensor of severe low oxygen levels in *Arabidopsis*, through its regulation of key hypoxia-response transcription factors. We found that plants lacking components of the N-end rule pathway constitutively express core hypoxia-response genes and are more tolerant of hypoxic stress. We identify the hypoxia-associated ethylene response factor group VII transcription factors of *Arabidopsis* as substrates of this pathway. Regulation of these proteins by the N-end rule pathway occurs through a characteristic conserved motif at the amino terminus initiating with Met-Cys. Enhanced stability of one of these proteins, HRE2, under low oxygen conditions improves hypoxia survival and reveals a molecular mechanism for oxygen sensing in plants via the evolutionarily conserved N-end rule pathway. SUB1A-1, a major determinant of submergence tolerance in rice<sup>3</sup>, was shown not to be a substrate for the N-end rule pathway despite containing the N-terminal motif, indicating that it is uncoupled from N-end rule pathway regulation, and that enhanced stability may relate to the superior tolerance of Sub1 rice varieties to multiple abiotic stresses<sup>4</sup>.

The N-end rule pathway of targeted proteolysis associates the fate of a protein substrate with the identity of its N terminus (the N-degron)<sup>5,6</sup>. The N-terminal residue is classified as stabilizing or destabilizing, depending on the fate of the protein. An N-degron containing a destabilizing residue is created through specific proteolytic cleavage, but can also be generated via successive enzymatic or chemical modifications to the N terminus; for example, arginylation by Arg-tRNA protein transferases (ATE)<sup>7–9</sup> (Supplementary Fig. 1). N-end rule pathway substrates containing destabilizing residues are targeted for proteasomal degradation via specific E3 ligases (also known as N-recognins), such as PROTEOLYSIS1 and PROTEOLYSIS6 (PRT1 and PRT6) in *Arabidopsis*, which accept substrates with hydrophobic and basic N termini, respectively<sup>8–10</sup>. Several substrates of the N-end rule pathway are important developmental regulators in mammals<sup>11</sup> but as yet no substrates have been identified in plants. Previously we showed a function of this pathway in abscisic acid (ABA) signalling through PRT6 and ATE<sup>12</sup>, and it has also been associated with leaf senescence and shoot and leaf development in *Arabidopsis*<sup>13,14</sup>. To understand N-end-rule-pathway-regulated gene expression we analysed the transcriptome of imbibed seed and seedlings of N-end rule pathway mutants *ate1 ate2*,

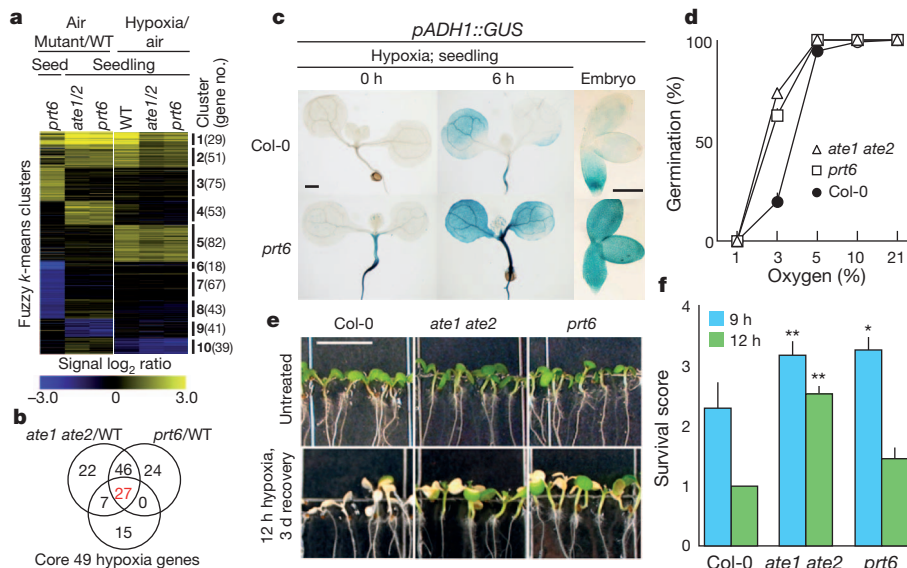
which lack ATE activity<sup>14</sup>, and *prt6* (Fig. 1a and Supplementary Table 1). This analysis revealed that genes important for anaerobic metabolism and survival of hypoxia, such as *ADH1*, *SUS4* and *PDC1*, were constitutively expressed at high levels in both mutants, in common with wild-type Col-0 plants under hypoxia (Supplementary Fig. 2). For example, 47 of the 135 differentially regulated mRNAs in the wild-type hypoxia-induced transcriptome were also upregulated in *prt6* seedlings grown under non-stress conditions (Supplementary Table 1; signal log<sub>2</sub> ratio ≥ 1, false discovery rate ≤ 0.01). The mRNAs upregulated in *prt6* and *ate1 ate2* mutants included over half of the core 49 mRNAs upregulated by hypoxia across seedling cell types<sup>15</sup> (Fig. 1b and Supplementary Fig. 2). Consistent with this observation, β-glucuronidase (GUS) expression driven by the promoter of *ADH1* (*pADH1::GUS*; ref. 16) was upregulated in wild-type seedlings subjected to hypoxia and ectopically expressed in mature embryos, roots and lower hypocotyls of *prt6* mutants (Fig. 1c and Supplementary Fig. 3). Constitutive expression of hypoxia-induced genes by N-end rule pathway mutant seedlings suggested that they would be resistant to hypoxic conditions. Imbibed seeds of both *prt6* and *ate1 ate2* mutants were able to germinate well under low oxygen (3%) compared to wild type (Fig. 1d), and mutant seedlings were more able to survive prolonged oxygen deprivation (Fig. 1e, f). The *ate1 ate2* double mutant showed greater resistance to hypoxia than *prt6*, indicating the existence of other as-yet-unidentified Arg-related E3 ligases, as previously postulated<sup>10,14</sup>.

Transcription factors of the five-member *Arabidopsis* ethylene response factor (ERF) group VII<sup>17</sup> have recently been shown to enhance plant responses to hypoxia or anoxia, including HYPOXIA RESPONSIVE1 and 2 (HRE1 and HRE2)<sup>18</sup> and RELATED TO AP2 2 (RAP2.2)<sup>19</sup>. Overexpression of RAP2.12 was also shown to induce expression of a *pADH1::LUCIFERASE* reporter gene<sup>20</sup>. This subfamily shows homology to the agronomically important rice ERFs SUBMERGENCE 1A, B and C (ref. 3) and SNORKEL 1 and 2 (ref. 21). *SUB1A-1* within the *SUBMERGENCE 1* (*SUB1*) locus (which also contains *SUB1B* and *SUB1C*) was shown to be a primary determinant of enhanced survival of rice plants under complete submergence<sup>3</sup>. With the exception of SUB1C, all contain the initiating motif Met-Cys (MC) at the N terminus, embedded within a longer consensus shared with most other group VII ERFs of *Arabidopsis* and rice, MCGGAI (Supplementary Fig. 4a).

Removal of N-terminal methionine by METHIONINE AMINO-PEPTIDASE (MAP) reveals the tertiary destabilizing residue cysteine in proteins initiating with MC, which targets substrates for degradation by the N-end rule pathway<sup>7,9,22</sup> (Supplementary Fig. 1). In mouse, N-end-rule-pathway-mediated degradation of the MC-motif-containing G-protein signalling components RGS4 and RGS5 is perturbed under hypoxia<sup>22,23</sup>. It was hypothesized that oxidation of cysteine at position 2 (C2) in these proteins under normoxia creates a secondary destabilizing

<sup>1</sup>Division of Plant and Crop Sciences, School of Biosciences and Centre for Plant Integrative Biology, University of Nottingham, Loughborough LE12 5RD, UK. <sup>2</sup>Center for Plant Cell Biology and Department of Botany and Plant Sciences, University of California, Riverside, California 92521, USA. <sup>3</sup>UPMC Univ Paris 06, UR5-EAC 7180 CNRS, Boîte courrier 156, 4 place Jussieu, F-75005 Paris, France. <sup>4</sup>Biological Chemistry Department, Rothamsted Research, Harpenden AL5 2JQ, UK.

\*These authors contributed equally to this work.



**Figure 1 | N-end rule mutants ectopically accumulate anaerobic response mRNAs and are more tolerant to hypoxia.** **a**, Expression data for differentially expressed genes comparing wild-type (Col-0) and mutants under air or hypoxia (2 h  $-O_2$ ). **b**, mRNAs upregulated in mutants overlap with 49 mRNAs induced across cell types by hypoxia in wild-type seedlings<sup>15</sup>. **c**, Spatial

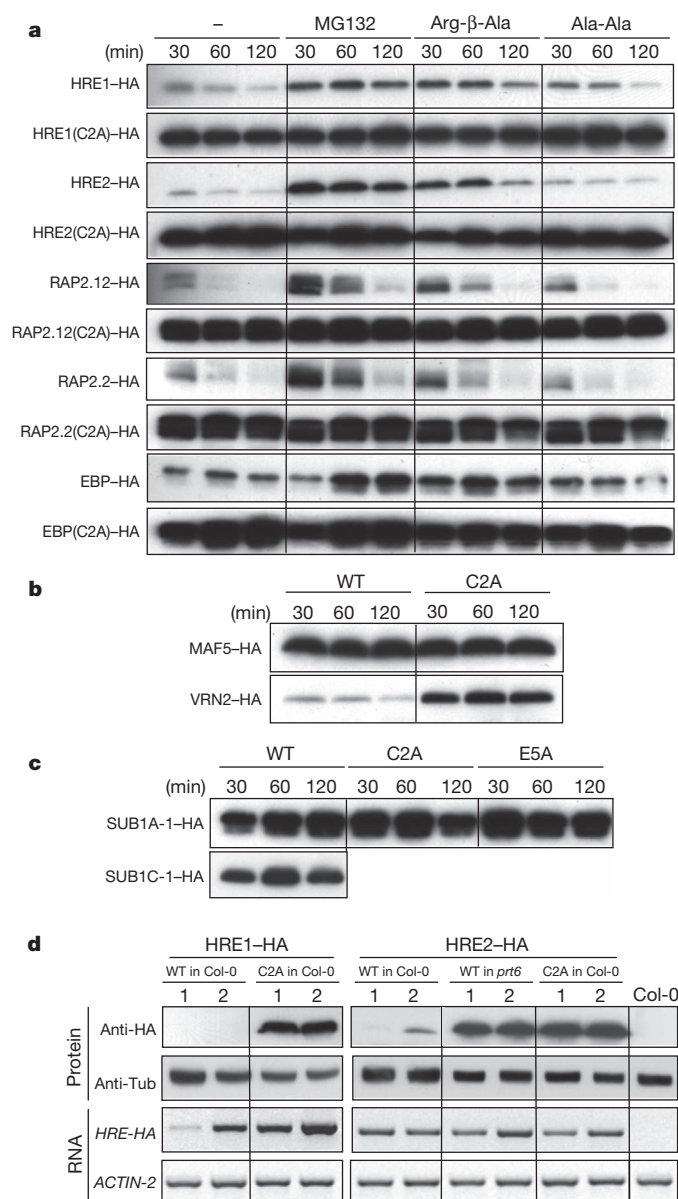
visualization of *ADH1* promoter activity. Scale bars: 100  $\mu$ m. **d**, Germination under reduced oxygen availability. **e**, Seedlings after 12 h of hypoxia and 3 d recovery. Scale bar: 0.6 cm. **f**, N-end rule pathway mutants are less sensitive to hypoxia stress. Data are mean of replicate experiments  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ .

residue allowing addition of arginine (R) to the N terminus by ATE, creating a primary destabilizing residue<sup>23</sup>. We investigated the possibility that all *Arabidopsis* group VII ERFs as well as rice SUB1A-1 are N-end rule pathway substrates. A heterologous rabbit reticulocyte lysate assay<sup>23</sup> was used to express haemagglutinin (HA)-tagged ERFs driven by a T7 promoter *in vitro*, because components of the N-end rule pathway (ATE, MAP and PRT6) are highly conserved in eukaryotes<sup>8</sup>, and it has been shown that wheat-germ lysate does not contain an active proteosomal system<sup>24</sup>. *Arabidopsis* group VII ERFs were short-lived, and their stability was enhanced by MG132 and the N-end rule pathway competitive dipeptide Arg- $\beta$ -Ala, but not by the non-competitive Ala-Ala dipeptide<sup>23</sup> (Fig. 2a). Mutation of C2 to alanine (C2A), which should remove the N-degron and stabilize proteins specifically with respect to the N-end rule pathway<sup>23</sup>, significantly enhanced stability *in vitro* of *Arabidopsis* ERFs, indicating that all group VII ERFs are potential substrates of the N-end rule pathway. *Arabidopsis* contains 206 proteins from gene models with MC at the N terminus; we used two of these—VERNALISATION 2 (VRN2) and MADS AFFECTING FLOWERING 5 (MAF5), which lack the extended N-terminal group VII ERF consensus (Supplementary Fig. 4b)—to test the specificity of this sequence. Whereas HA-tagged VRN2 (VRN2-HA) was degraded in this system, and stabilized by the introduction of a C2A mutation (VRN2(C2A)-HA), MAF5-HA and MAF5(C2A)-HA were both stable (Fig. 2b), indicating that not all *Arabidopsis* MC proteins are N-end rule pathway substrates. This is not surprising as it has previously been shown that optimal positioning of a downstream lysine for ubiquitination is also a key determinant of the quality of an N degron<sup>8,9,25</sup>. SUB1A-1 was resistant to degradation (Fig. 2c). As the N-terminal sequence of SUB1A-1 differs at position 5 (E rather than A, Supplementary Fig. 4a), we analysed a mutant version that replaced this amino acid to reconstitute the consensus group VII sequence (SUB1A-1(E5A)-HA). SUB1A-1(E5A)-HA was also stable *in vitro* (Fig. 2c), indicating that degradation of this protein is uncoupled from the N-end rule pathway. As expected, the rice protein SUB1C-1-HA, lacking an MC N terminus, was long lived *in vitro* (Fig. 2c).

To confirm the activity of the N-end rule pathway towards specific MC-containing substrates in plants, we analysed the *in vivo* longevity of the ERF proteins HRE1 and HRE2 (Fig. 2d). We expressed either

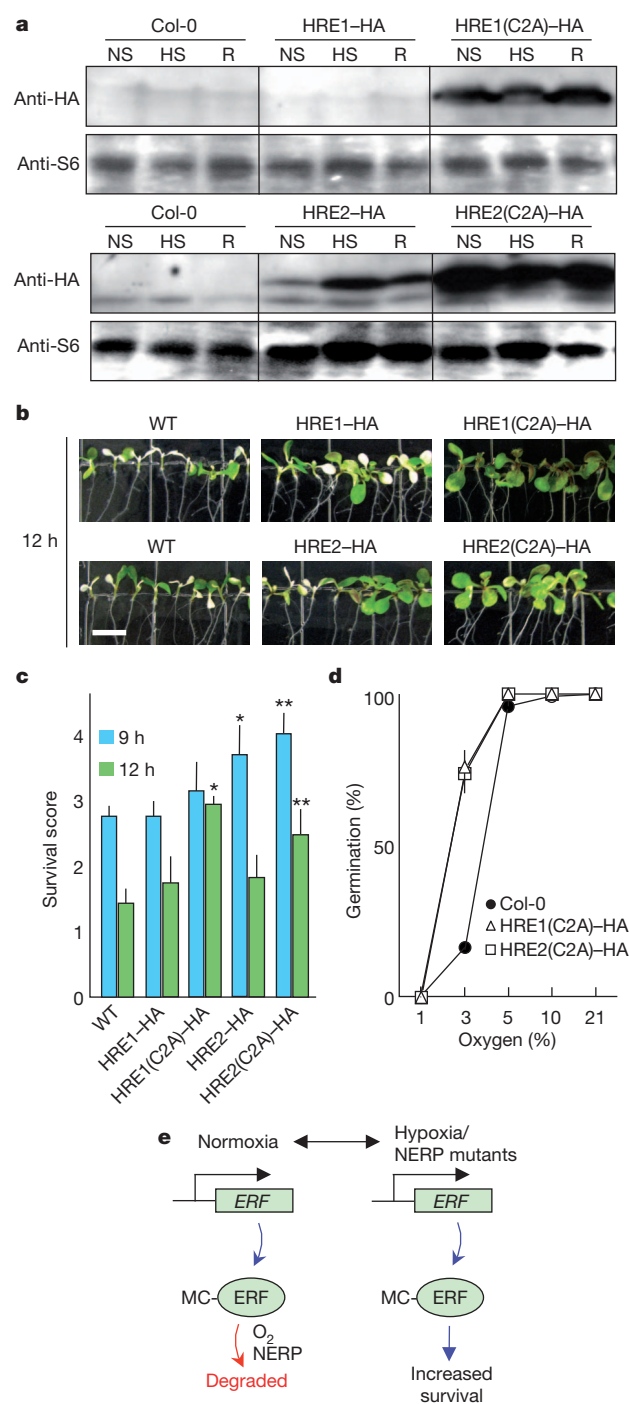
wild-type or mutant (HRE1(C2A), HRE2(C2A)) HA-tagged versions of these proteins ectopically using the CaMV35S promoter in *Arabidopsis*. In wild-type plants, only the mutant C2A proteins could be detected at high levels, despite detectable expression of corresponding mRNAs, indicating that wild-type versions are N-end rule pathway substrates *in vivo*. HRE2-HA expressed in the *prt6* mutant was stable, linking its degradation directly to PRT6. To assess whether oxygen regulates the stability of HRE proteins, we analysed the accumulation of HRE-HA proteins in wild-type plants expressing HRE1-HA, HRE1(C2A)-HA, HRE2-HA and HRE2(C2A)-HA under normal and low oxygen conditions (Fig. 3a). After transfer of seedlings to hypoxic conditions we observed elevation of HRE2-HA within 2 h, but could not detect HRE1-HA (Fig. 3a and Supplementary Fig. 5a, b). HRE2-HA became destabilized again upon return to normoxic conditions (Fig. 3a). Both seeds and seedlings ectopically expressing stable C2A versions of HRE1 and HRE2 had increased tolerance to extended periods of oxygen deprivation (Fig. 3b–d and Supplementary Fig. 5c).

These data demonstrate that *Arabidopsis* ERF group VII transcription factors are substrates of the N-end rule pathway, and function to sense molecular oxygen, most likely through oxidation of the tertiary destabilizing residue cysteine. Stabilization of these proteins under hypoxic conditions leads to increased survival under low oxygen stress (Fig. 3e). It is currently unclear whether oxidation occurs through a chemical or enzymatic mechanism, although cysteine is readily oxidized chemically<sup>26</sup>. It is also unclear whether oxidation is related directly to molecular oxygen, or if indirect cellular changes associated with oxygen availability (such as alterations in cytosolic pH<sup>27</sup> and specific metabolites or transient accumulation of reactive oxygen species<sup>4</sup>) might trigger cysteine oxidation. SUB1A-1 may provide enhanced responsiveness to submergence and drought in rice in part due to the fact that it is not a substrate of the N-end rule pathway. By contrast, the condition-dependent destabilization of group VII ERFs in *Arabidopsis* could require oxygen levels to decline below some threshold before these factors can activate anaerobic gene transcription. It is probable that SUB1A-1 evades the N-end rule pathway due to the absence of an optimally positioned lysine downstream of the N degron, as substrate quality is determined combinatorially by an N degron destabilizing residue and downstream lysine position<sup>8,9,25</sup>. Alternatively, differences in protein tertiary structure may preclude



**Figure 2 | Group VII ERF transcription factors are substrates for the N-end rule pathway *in vitro* and *in vivo*.** **a**, Western blot analysis of *in vitro* stability of HA-tagged wild-type and C2A variants of *Arabidopsis* group VII ERFs in the absence or presence of MG132, N-end rule pathway competitive dipeptide (Arg-β-Ala) or non-competitive dipeptide (Ala-Ala). **b**, *In vitro* stability of wild-type and C2A VRN2-HA and MAF5-HA. **c**, *In vitro* stability of HA-tagged rice ERFs. **d**, *In vivo* protein stability and RNA expression levels of wild-type and C2A variants of HRE1-HA and HRE2-HA ectopically expressed in *Arabidopsis*, shown for two independent transformed lines (1 and 2).

N-terminus accessibility. SUB1A-1 was also recently shown to mediate crosstalk between submergence and drought tolerance in rice by augmenting ABA responsiveness<sup>4</sup>, suggesting a link between drought tolerance and the previously identified function of the N-end rule pathway in removing responsiveness to ABA<sup>12</sup>. Targeted degradation of proteins by the N-end rule pathway was identified as a homeostatic mechanism in mammalian systems<sup>22,23,28</sup>, for example in the control of hypoxia-related expression of RGS4 (ref. 28) and RGS5 (ref. 23). It is fascinating that the N-end rule pathway carries out the same functionality in relation to low oxygen stress in plants, but taking as substrates members of a plant-specific transcription factor family. This highlights evolutionary conservation of the mechanism of oxygen perception across kingdoms using the N-end rule pathway independent of the targets. Our confirmation of *in vivo* function of two members of ERF



**Figure 3 | HRE proteins are stabilized under low oxygen levels and confer hypoxia tolerance.** **a**, *In vivo* stability of wild-type and C2A HRE1-HA, HRE2-HA (anti-HA) or S6 (ribosomal protein S6) control (anti-S6). HS, 2 h hypoxia; NS, no stress; R, following 1 h recovery from stress. **b**, Seedlings expressing wild-type or C2A HRE1-HA and HRE2-HA after 12 h hypoxic stress and 3 d of recovery. Scale bar: 0.6 cm. **c**, Seedling survival for wild-type or C2A HRE1-HA and HRE2-HA after 9 h or 12 h hypoxic stress. Data are mean of replicate experiments  $\pm$  s.d. \* $P < 0.05$ ; \*\* $P < 0.01$ . **d**, Germination under reduced oxygen availability. **e**, Model explaining N-end-rule-pathway-mediated oxygen-dependent turnover of group VII ERFs in *Arabidopsis*.

group VII provides direct evidence for the control of HRE2 by oxygen and the N-end rule pathway and indirect evidence that HRE1 is also an N-end rule pathway substrate *in vivo*. We demonstrate that all members of *Arabidopsis* group VII ERFs are N-end rule pathway substrates *in vitro*, and thus it is possible that all members orchestrate N-end-rule-pathway-controlled, hypoxia-related functions. Identification and

manipulation of N-end rule pathway substrates will therefore be a key target for both conventional breeding and biotechnological approaches in relation to manipulation of plant responses to abiotic stress.

## METHODS SUMMARY

**Protein stability analyses.** Full-length cDNAs were amplified by polymerase chain reaction (PCR) from either *Arabidopsis thaliana* or *Oryza sativa* L. (cv. M202(Sub1)). N-terminal mutations were introduced using the forward primer (Supplementary Table 2). For *in vitro* assays, cDNAs were cloned into a modified version of the pTNT vector (Promega) to produce C-terminal HA fusions. Stability assays were performed using the TNT T7 Coupled Reticulocyte Lysate system (Promega), essentially as described previously<sup>23</sup>. For *in vivo* analysis of HRE–HA proteins, cDNAs were cloned into pE2c, mobilized into pB2GW7 and transformed into *Arabidopsis* using the floral dip method. To assess relative protein stability, equal amounts of total protein extracted from 7-day-old T<sub>3</sub> homozygous seedlings were analysed by western blot, and cDNA synthesized from total RNA was used as a template for semi-quantitative PCR.

**Gene expression analyses.** For microarray analysis, total RNA extracted from seeds<sup>12</sup> or seedlings<sup>15</sup> was hybridized against the *Arabidopsis* ATH1 genome array (Affymetrix). Differentially expressed genes were clustered as described previously<sup>15</sup>. *pADH::GUS*<sup>16</sup> was crossed to *prt6-1* and homozygous seeds or seedlings were analysed for GUS activity before and after submergence for the times indicated.

**Low O<sub>2</sub> phenotypic analyses.** To assess germination (scored as radicle emergence), imbibed seeds were incubated for 7-days in chambers flushed with varying O<sub>2</sub> tensions<sup>29</sup>. For 7-day-old seedling survival, O<sub>2</sub> deprivation was achieved by bubbling 99.995% argon through water into chambers under positive pressure, before recovering in air for 3 days and scoring of plants (*n* = 15) per plate that were non-damaged, damaged or dead (scored 5, 3 and 1, respectively)<sup>15</sup>. The same argon chambers were used to treat seedlings for the times indicated before protein extraction for western blot analysis.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 13 June; accepted 5 September 2011.

Published online 23 October 2011.

1. Bailey-Serres, J. & Voesenek, L. A. C. J. Flooding stress: Acclimations and genetic diversity. *Annu. Rev. Plant Biol.* **59**, 313–339 (2008).
2. Kaelin, W. G. & Ratcliffe, P. J. Oxygen sensing by metazoans: The central role of the HIF hydroxylase pathway. *Mol. Cell* **30**, 393–402 (2008).
3. Xu, K. *et al.* *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**, 705–708 (2006).
4. Fukao, T., Yeung, E. & Bailey-Serres, J. The submergence tolerance regulator SUB1A mediates crosstalk between submergence and drought tolerance in rice. *Plant Cell* **23**, 412–427 (2011).
5. Bachmair, A., Finley, D. & Varshavsky, A. *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186 (1986).
6. Varshavsky, A. Regulated protein degradation. *Trends Biochem. Sci.* **30**, 283–286 (2005).
7. Kwon, Y. T. *et al.* An essential role of N-terminal arginylation in cardiovascular development. *Science* **297**, 96–99 (2002).
8. Graciet, E., Mesiti, F. & Wellmer, F. Structure and evolutionary conservation of the plant N-end rule pathway. *Plant J.* **61**, 741–751 (2010).
9. Graciet, E. & Wellmer, F. The plant N-end rule pathway: structure and functions. *Trends Plant Sci.* **15**, 447–453 (2010).
10. Garzón, M. *et al.* PRT6/At5g02310 encodes an *Arabidopsis* ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus. *FEBS Lett.* **581**, 3189–3196 (2007).

11. Tasaki, T. & Kwon, Y. T. The mammalian N-end rule pathway: new insights into its components and physiological roles. *Trends Biochem. Sci.* **32**, 520–528 (2007).
12. Holman, T. J. *et al.* The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **106**, 4549–4554 (2009).
13. Yoshida, S., Ito, M., Callis, J., Nishida, I. & Watanabe, A. A delayed leaf senescence mutant is defective in arginyl-tRNA: protein arginyltransferase, a component of the N-end rule pathway in *Arabidopsis*. *Plant J.* **32**, 129–137 (2002).
14. Graciet, E. *et al.* The N-end rule pathway controls multiple functions during *Arabidopsis* shoot and leaf development. *Proc. Natl Acad. Sci. USA* **106**, 13618–13623 (2009).
15. Mustroph, A. *et al.* Profiling translatores of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **106**, 18843–18848 (2009).
16. Chung, H. J. & Ferl, R. J. *Arabidopsis* alcohol dehydrogenase expression in both shoots and roots is conditioned by root growth environment. *Plant Physiol.* **121**, 429–436 (1999).
17. Nakano, T., Suzuki, K., Fujimura, T. & Shinshi, H. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* **140**, 411–432 (2006).
18. Licausi, F. *et al.* HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana*. *Plant J.* **62**, 302–315 (2010).
19. Hinz, M. *et al.* *Arabidopsis* RAP2.2: An ethylene response transcription factor that is important for hypoxia survival. *Plant Physiol.* **153**, 757–772 (2010).
20. Papdi, C. *et al.* Functional identification of *Arabidopsis* stress regulatory genes using the controlled cDNA overexpression system. *Plant Physiol.* **147**, 528–542 (2008).
21. Hattori, Y. *et al.* The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature* **460**, 1026–1030 (2009).
22. Hu, R. G. *et al.* The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature* **437**, 981–986 (2005).
23. Lee, M. J. *et al.* RGS4 and RGS5 are *in vivo* substrates of the N-end rule pathway. *Proc. Natl Acad. Sci. USA* **102**, 15030–15035 (2005).
24. Takahashi, H. *et al.* A simple and high-sensitivity method for analysis of ubiquitination and polyubiquitination based on wheat cell-free protein synthesis. *BMC Plant Biol.* **9**, 39 (2009).
25. Suzuki, T. & Varshavsky, A. Degradation signals in the lysine-asparagine sequence space. *EMBO J.* **18**, 6017–6026 (1999).
26. Leonard, S. E. & Carroll, K. S. Chemical ‘omics’ approaches for understanding protein cysteine oxidation in biology. *Curr. Opin. Chem. Biol.* **15**, 88–102 (2011).
27. Felle, H. H. pH regulation in anoxic plants. *Ann. Bot.* **96**, 519–532 (2005).
28. Hu, R. G., Wang, H. Q., Xia, Z. X. & Varshavsky, A. The N-end rule pathway is a sensor of heme. *Proc. Natl Acad. Sci. USA* **105**, 76–81 (2008).
29. Côme, D. & Tissaoui, T. Induction d’une dormance embryonnaire secondaire chez le pommier (*Pirus malus* L.) par des atmosphères très appauvries en oxygène. *Compt. Rendus Acad. Sci.* **266**, 477–479 (1968).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** M.J.H., D.J.G., S.G. and C.S.C. were supported by BBSRC grant BB/G010595/1; G.W.B. by a Marie Curie International Incoming Fellowship; N.M.I. by a MARA PhD fellowship from the Malaysian government; S.C.L., T.F. and J.B.-S. by grants NSF IOS-0750811 and NIFA 2008-35100-04528. We thank S. Liddell. Rothamsted Research receives grant-aided support from the BBSRC.

**Author Contributions** D.J.G., M.J.H., J.B.-S., F.C. and F.L.T. conceived and designed experiments. D.J.G., S.C.L., N.M.I., S.G., C.S.C., G.W.B., T.F. and F.C. performed the experiments. D.J.G., S.C.L., N.M.I., S.G., C.S.C., G.W.B., T.F., F.C., M.J.H., J.B.-S. and F.L.T. analysed the data. M.J.H., D.J.G. and J.B.-S. wrote the manuscript.

**Author Information** The microarray data reported in this paper are deposited in Gene Expression Omnibus under accession number GSE29941 and are also tabulated in Supplementary Information. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.J.H. ([michael.holdsworth@nottingham.ac.uk](mailto:michael.holdsworth@nottingham.ac.uk)) or J.B.-S. ([serres@ucr.edu](mailto:serres@ucr.edu)).

## METHODS

**Growth and analysis of plant material.** *Arabidopsis thaliana* seeds were obtained from NASC, except for transgenics containing *pADH::GUS* (ref. 16) (a gift from R. Ferl). Columbia-0 (Col-0) was the wild type for all analyses. *prt6-1*, *prt6-5* and *ate1 ate2* mutants were described previously<sup>12,14</sup>. For the generation of transgenic *Arabidopsis* and *in vivo* protein assays, plants were grown vertically on half MS media for 7 days at 22 °C in 150 µmol m<sup>-2</sup> s<sup>-1</sup> constant light and transferred to soil after 2 weeks if required. For analysis of seedling O<sub>2</sub> deprivation survival and protein analysis, plants were grown vertically on MS medium (0.43% (w/v) MS salts, 1% (w/v) Suc and 0.4% (w/v) phytagel, pH 5.75) at 23 °C with a 16-h-day (50 µmol m<sup>-2</sup> s<sup>-1</sup>) and 8-h-night cycle for 7 d. The rice (*Oryza sativa* L.) *SUB1* introgression line cv. M202(*Sub1*) was grown and submerged before cDNA isolation as described previously<sup>3</sup>. All plant experiments were carried out at least three times.

**Analysis of oxygen deprivation response in seeds and seedlings.** Seven-day-old *Arabidopsis* seedlings were subjected for specified durations to non-stress (NS) or hypoxia stress (HS) treatments, or subjected to hypoxia stress and returned to ambient air (re-oxygenation; R). For seedling survival, 15 Col-0 and 15 mutant seedlings were grown side by side (3 replicates). Treatments commenced at the end of the 16-h light cycle in open (NS) or sealed (HS) chambers. For HS, 99.995% argon gas was bubbled through water and into the chamber while air was expelled by positive pressure<sup>30</sup>. After treatment, the 15 seedlings per genotype per plate were scored as non-damaged, damaged and dead (scored 5, 3 and 1, respectively) compared to wild-type plants grown on the same plate and results analysed using the students *t*-test, as described previously<sup>31</sup>, or seedlings were frozen under liquid nitrogen within 3 min of release before protein extraction.

Germination of *Arabidopsis* seeds (3–4 replicates of *n* = 60–100; scored on day 7 as radicle emergence) was performed at 22 °C under constant light in various oxygen tensions achieved through mixing N<sub>2</sub> and air via capillary tubes according to the apparatus described previously<sup>29</sup>.

Wild-type plants carrying the *pADH::GUS* transgene<sup>16</sup> were crossed to *prt6-1* plants and homozygous *prt6-1 pADH::GUS* individuals were identified in the F<sub>2</sub> population. Seven-day-old seedlings were submerged in degassed water in the dark to induce hypoxia for the times indicated. Embryos were dissected 6 h after being imbibed. Seedlings and embryos were assayed for GUS activity and imaged following standard methods<sup>32</sup>.

**Construction of transgenic plants and protein and RNA extractions.** To generate C-terminally HA-tagged ERF fusions of *HRE1* (At1g72360) and *HRE2* (At2g47520) driven by the 35SCaMV promoter, full-length cDNAs amplified from *Arabidopsis* total seedling cDNA were first ligated into the Entry vector pE2c and then mobilized into the Destination binary vector pB2GW7, as described previously<sup>33</sup>. N-terminal mutations were incorporated by changing the forward primer sequences accordingly (Supplementary Table 2). Transformation into *Agrobacterium tumefaciens* (strain GV3101 pMP90) and *Arabidopsis thaliana* was performed according to established protocols<sup>34</sup>. Proteins were extracted from 7-day-old homozygous T<sub>3</sub> seedlings as described<sup>35</sup>. Extracts were quantified using the Bio-Rad DC assay and subjected to anti-HA immunoblot analysis. For semi-quantitative RT-PCR, RNA was extracted using an RNEasy plant mini kit (Qiagen) and converted to cDNA using Superscript III Reverse transcriptase (Invitrogen). PCRs were performed with transgene-specific primers (gene-specific forward, HA-tag reverse) and *ACTIN-2* was amplified for use as a loading control (Supplementary Table 2).

**In vitro analysis of protein stability.** To generate *Arabidopsis* and rice protein-HA fusions driven by the T7 promoter, cDNAs were PCR amplified from *Arabidopsis* total cDNA or submerged rice cDNA (M202(*Sub1*)), as described<sup>3</sup>, and ligated into a modified version of the pTNT (Invitrogen) expression vector (pTNT3xHA). N-terminal mutations were incorporated by changing the forward primer sequences accordingly (Supplementary Table 2).

Proteins were expressed *in vitro* using the TNT T7 Coupled Reticulocyte Lysate system (Promega) according to manufacturer's guidelines, using 500 ng plasmid template. Where appropriate, 100 µM MG132 or 1 mM dipeptides (Arg-β-Ala or Ala-Ala; Sigma-Aldrich) and 150 nM Bestatin (Sigma-Aldrich) were added. Reactions were incubated at 30 °C, and samples were taken at indicated time points before mixing with protein loading dye to terminate protein synthesis. Equal amounts of each reaction were subjected to anti-HA immunoblot analysis. All blots were checked for equal loading by Ponceau staining.

**Immunoblotting.** Proteins resolved by SDS-PAGE were transferred to PVDF using a MiniTrans-Blot electrophoretic transfer cell (Bio-Rad). Membranes were probed with primary antibodies at the following titres: anti-HA (Sigma-Aldrich), 1:1,000; anti-α-tubulin (Sigma-Aldrich), 1:5,000; anti-ribosomal protein S6 (ref. 36), (1:5,000). HRP-conjugated anti-mouse secondary antibody (Santa Cruz) was used at a titre of 1:10,000. Immunoblots were developed to film using ECL western blotting substrate (Pierce).

**Alignment of MC-ERF proteins from *Arabidopsis* and rice.** Rice and *Arabidopsis* ERF proteins starting with the sequence MC were aligned and phylogenetic relationships observed using CLUSTALW<sup>37</sup>.

**Microarray hybridization and data analyses.** Total RNA extracted from seeds or seedlings was assessed for quality using the Agilent 2100 Bioanalyser with the RNA 6000 Nano reagent kit. Biotin-labelled cRNA was synthesized using the Affymetrix 3' IVT Express Labelling kit and hybridized against the *Arabidopsis* ATH1 genome array (GeneChip System, Affymetrix). CEL file data were processed to estimate the abundance of each expressed mRNA in two (seedling) or three (imbibed seed) biological replicate samples as described previously<sup>15</sup>. The microarray experiments reported here are described following MIAME guidelines and are deposited in GEO under the accession number GSE29941.

The differentially expressed genes were further analysed by use of fuzzy *k*-means clustering with the FANNY function from the Cluster package in R, as described<sup>15</sup>. The resulting gene-to-cluster assignments are given in Supplementary Table 1 and were visualized with the TIGR MEV program. Each gene cluster was evaluated for enrichment of specific gene functions (Gene Ontology (GO)) as described previously<sup>38</sup> using *Arabidopsis* gene-to-GO mappings from TAIR (<http://geneontology.org>; downloaded 17 May 2011).

30. Branco-Price, C., Kaiser, K. A., Jang, C. J. H., Larive, C. K. & Bailey-Serres, J. Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. *Plant J.* **56**, 743–755 (2008).
31. Mustroph, A. *et al.* Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. *Plant Physiol.* **152**, 1484–1500 (2010).
32. Lucas, M. *et al.* SHORT-ROOT regulates primary, lateral, and adventitious root development in *Arabidopsis*. *Plant Physiol.* **155**, 384–398 (2011).
33. Dubin, M. J., Bowler, C. & Benvenuto, G. A modified Gateway cloning strategy for overexpressing tagged proteins in plants. *Plant Methods* **4**, 3 (2008).
34. Swarup, R. *et al.* Root gravitropism requires lateral root cap and epidermal cells for transport and response to a mobile auxin signal. *Nature Cell Biol.* **7**, 1057–1065 (2005).
35. Martinez-Garcia, J. F., Monte, E. & Quail, P. H. A simple, rapid and quantitative method for preparing *Arabidopsis* protein extracts for immunoblot analysis. *Plant J.* **20**, 251–257 (1999).
36. Williams, A. J., Werner-Fraczek, J., Chang, I. F. & Bailey-Serres, J. Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. *Plant Physiol.* **132**, 2086–2097 (2003).
37. Thompson, J. D., Higgins, D. G. & Gibson, T. J. Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
38. Horan, K. *et al.* Annotating genes of known and unknown function by large-scale coexpression analysis. *Plant Physiol.* **147**, 41–57 (2008).

# Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization

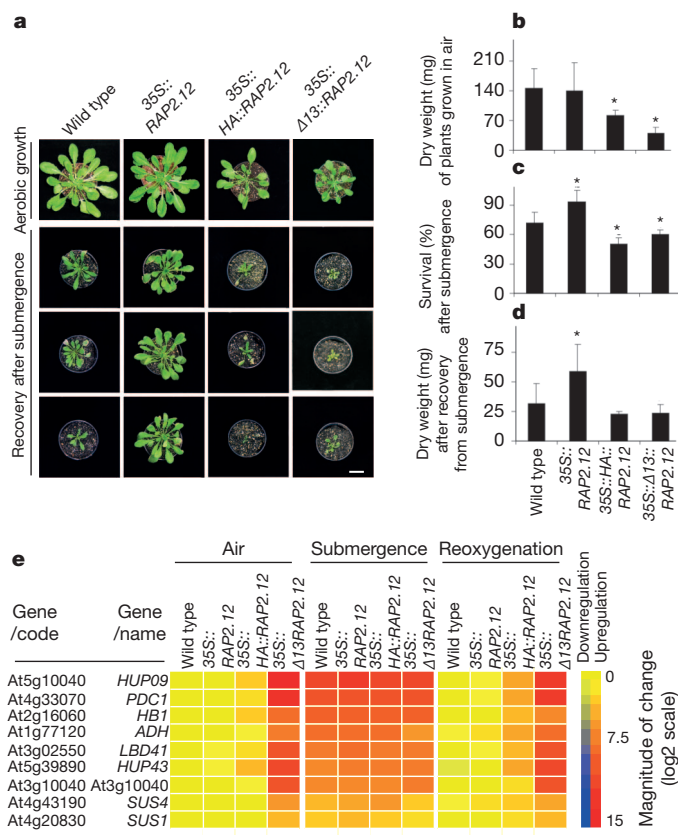
Francesco Licausi<sup>1,2</sup>, Monika Kosmacz<sup>1</sup>, Daan A. Weits<sup>1</sup>, Beatrice Giuntoli<sup>2</sup>, Federico M. Giorgi<sup>1</sup>, Laurentius A. C. J. Voesenek<sup>3,4</sup>, Pierdomenico Perata<sup>2</sup> & Joost T. van Dongen<sup>1</sup>

The majority of eukaryotic organisms rely on molecular oxygen for respiratory energy production<sup>1</sup>. When the supply of oxygen is compromised, a variety of acclimation responses are activated to reduce the detrimental effects of energy depletion<sup>2–4</sup>. Various oxygen-sensing mechanisms have been described that are thought to trigger these responses<sup>5–9</sup>, but they each seem to be kingdom specific and no sensing mechanism has been identified in plants until now. Here we show that one branch of the ubiquitin-dependent N-end rule pathway for protein degradation, which is active in both mammals and plants<sup>10,11</sup>, functions as an oxygen-sensing mechanism in *Arabidopsis thaliana*. We identified a conserved amino-terminal amino acid sequence of the ethylene response factor (ERF)-transcription factor RAP2.12 to be dedicated to an oxygen-dependent sequence of post-translational modifications, which ultimately lead to degradation of RAP2.12 under aerobic conditions. When the oxygen concentration is low—as during flooding—RAP2.12 is released from the plasma membrane and accumulates in the nucleus to activate gene expression for hypoxia acclimation. Our discovery of an oxygen-sensing mechanism opens up new possibilities for improving flooding tolerance in crops.

Tolerance to submergence and low oxygen availability (hypoxia) have been considered to be influenced by different members of subgroup VII of the ERF transcription factor family in *Arabidopsis* (RAP2.12 (ref. 12), RAP2.2 (ref. 13); HRE1 and HRE2 (ref. 14)) and rice (SUB1 (ref. 15), SK1 and SK2 (ref. 16)). Here, we reveal the mechanism by which molecular oxygen acts upon RAP2.12 (At1g53910) to trigger molecular acclimation responses. RAP2.12 is highly homologous to RAP2.2 and is widely conserved in higher plants (Supplementary Fig. 1). It is constitutively expressed throughout the entire plant (Supplementary Fig. 2) and further upregulated in leaves upon hypoxia, but not by the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) (Supplementary Fig. 3). RAP2.12 positively regulates gene transcription *in planta* via a conserved carboxy-terminal motif (Supplementary Fig. 4). Constitutive overexpression of RAP2.12 (35S::RAP2.12) did not significantly affect the phenotype of *Arabidopsis* plants when grown aerobically (Fig. 1a, b). However, submergence tolerance of independently transformed 35S::RAP2.12 plants increased with respect to the wild-type control, as demonstrated by the increased number and dry weight of plants that recovered from submergence (Fig. 1a, c, d), which can be explained by the faster and stronger induction of hypoxia-responsive genes during the flooding treatment in 35S::RAP2.12 plants (Supplementary Fig. 5). Interestingly, different flooding-tolerance strategies in two wild *Rumex* species correlated with the differential induction of *ERF1*, which is the orthologue of RAP2.12 (Supplementary Fig. 6). In contrast, constitutive expression of RAP2.12 with a haemagglutinin (HA)-peptide tag at its N terminus (35S::HA::RAP2.12) resulted in a reduction of plant growth in air (Fig. 1a, b). Concomitantly, tolerance to submergence decreased as compared to the wild type (Fig. 1c). Similar results were observed when

a version of RAP2.12 was expressed from which the first 13 amino acid residues were deleted (35S::Δ13RAP2.12). It thus seemed that manipulating the N-terminal amino acid sequence obstructed the regulative function of RAP2.12 already under aerobic conditions, thereby reducing the vigour and stress tolerance of the plants.

To understand the impact of the N-terminal modifications on the activity of RAP2.12, we investigated which genes are expressed under the control of RAP2.12. We found that under aerobic conditions



**Figure 1 | The transcription factor RAP2.12 regulates hypoxia tolerance of plants.** **a**, The effect of overexpression of RAP2.12, HA::RAP2.12 or Δ13RAP2.12 on plant growth in air, or after submergence. Scale bar, 2 cm. **b**, Dry weight of 7-week-old rosette leaves from air-grown plants ( $n = 20$ ). **c**, Percentage of plants surviving flooding-induced hypoxia ( $n = 4$ ). **d**, Dry weight of rosette leaves from surviving plants, 2 weeks after the flooding treatment ( $n = 20$ ). **e**, Differential expression of hypoxia-responsive genes (reference: wild type in air). Numeric expression values are shown in Supplementary Table 1. Data are presented as mean  $\pm$  s.d. \* $P < 0.05$ , one-way ANOVA.

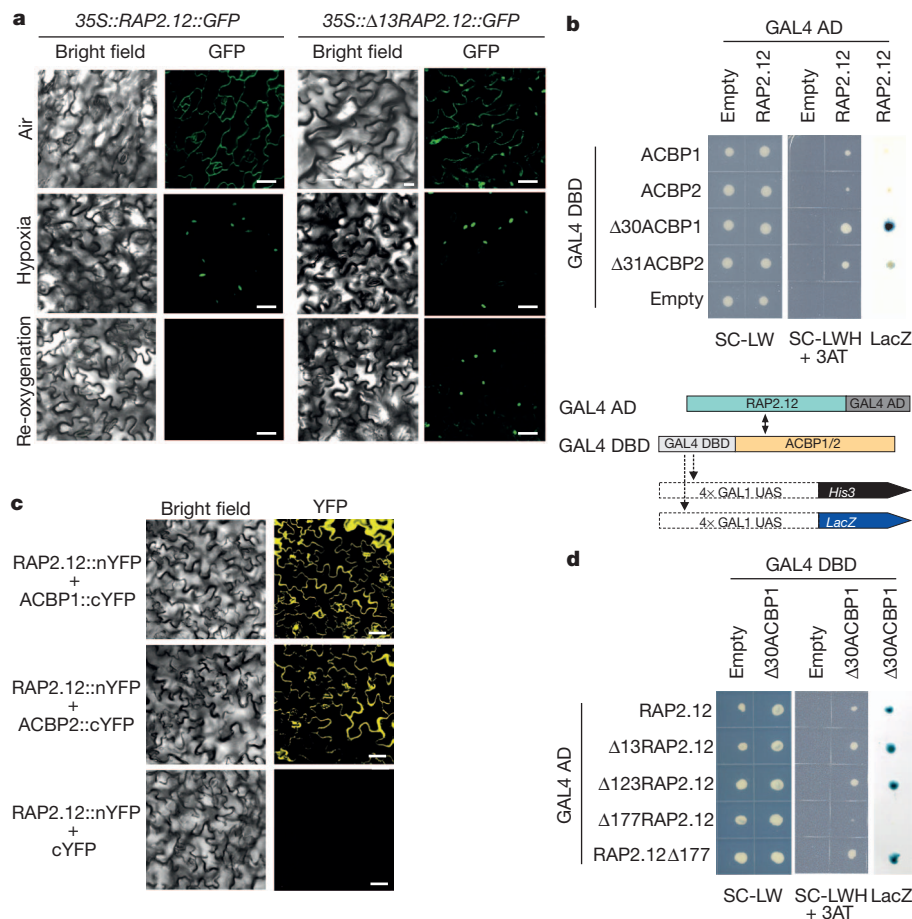
<sup>1</sup>Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476, Potsdam-Golm, Germany. <sup>2</sup>PlantLab, Institute of Life Sciences, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56127 Pisa, Italy. <sup>3</sup>Plant Ecophysiology, Institute of Environmental Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands. <sup>4</sup>Centre for Biosystems Genomics, 6708 PB Wageningen, the Netherlands.

35S::RAP2.12 plants exhibited a slight increase in the expression of hypoxia marker genes, whereas during flooding the expression of these hypoxia marker genes was more strongly upregulated in plants overexpressing RAP2.12 as compared to wild-type plants during flooding (Fig. 1e and Supplementary Table 1). During re-oxygenation, the expression of the hypoxia marker genes was rapidly downregulated in both wild-type plants and 35S::RAP2.12, whereas in 35S::HA::RAP2.12 and 35S::Δ13RAP2.12 the level of expression remained high, as before the flooding treatment. The correlation between this expression pattern of hypoxia response genes and the reduced growth and recovery after flooding that is observed for plants overexpressing HA::RAP2.12 and Δ13RAP2.12 indicates that proper upregulation of the hypoxia response genes during flooding as well as downregulation of these genes during recovery from flooding are both required for optimal plant acclimation. The activation of hypoxic gene expression by RAP2.12 was further confirmed by the observation that RAP2.12 induced a luciferase (*Luc*) reporter gene when its promoter contained the motif ATCTA (Supplementary Fig. 7), which was previously identified as a hypoxia-responsive element in plants<sup>17</sup>. On the other hand, the relatively small effect of 35S::RAP2.12 on gene expression under aerobic conditions (Fig. 1e) indicated that an additional regulatory mechanism reliant on sensing of low oxygen concentrations is needed to induce hypoxic gene expression. Interestingly, this requirement was abolished when the N terminus of RAP2.12 was modified either by fusing the HA-peptide tag to the protein (35S::HA::RAP2.12 in Fig. 1e), or by deleting its first conserved amino acid residues (35S::Δ13RAP2.12 in Fig. 1e and Supplementary Fig. 8), indicating that the N terminus of RAP2.12

has an important role in the regulation of the oxygen-dependent activation of the transcription factor.

Further comparative analysis of a full-genome expression profile of the hypoxic response in wild-type plants and the differential regulation of genes by expressing the HA::RAP2.12 construct under aerobic conditions revealed that the genes that were most strongly up- or downregulated by HA::RAP2.12 were also differentially expressed under hypoxia (Supplementary Fig. 9 and Supplementary Table 2). Similarly, the silencing of RAP2.12 and its closest homologue RAP2.2 using an artificial microRNA approach reduced the induction of hypoxic gene expression by low oxygen (Supplementary Fig. 10 and Supplementary Tables 3 and 4). Given that the messenger RNA stability of RAP2.12 was not affected by the additional nucleotides encoding the N-terminal peptide tag (Supplementary Fig. 11), we concluded that post-translational modifications of the N-terminal amino acid residues of RAP2.12 are involved in regulating the activity of this transcription factor, which is required to induce hypoxia core-response genes.

We further investigated the role of the N-terminal amino acid residues by determining the subcellular localization of RAP2.12 fused to green fluorescent protein (GFP). Under aerobic conditions, the fusion protein localized to the plasma membrane; however, upon hypoxia it accumulated in the nucleus (Fig. 2a and Supplementary Fig. 12). Remarkably, upon re-oxygenation the RAP2.12::GFP signal fully disappeared within 1 h (Fig. 2a). After deleting the conserved N-terminal amino acid residues of RAP2.12 (35S::Δ13RAP2.12::GFP), the transcription factor was observed in both the cell membrane and the nucleus under aerobic conditions (Fig. 2a). However, under hypoxia, the



**Figure 2 | RAP2.12 is membrane localized and re-localizes in the nucleus upon hypoxia.** **a**, Subcellular localization of stably transformed GFP-fused RAP2.12 and Δ13RAP2.12. Localization controls are shown in Supplementary Fig. 12. **b**, Yeast two-hybrid analysis showing interaction between RAP2.12 and ACBP1 and ACBP2. **c**, Bimolecular fluorescence complementation of YFP

confirming interaction between RAP2.12 and ACBP1 and ACBP2. **d**, Yeast two-hybrid analysis between various truncated RAP2.12 proteins and Δ30ACBP1. Names of genes are explained in a pictogram shown in Supplementary Fig. 16. AD, activation domain; DBD, DNA-binding domain; UAS, upstream activator sequence. Scale bars, 10 μm.

membrane association disappeared, the protein accumulated only in the nucleus and remained there even after re-oxygenation (35S::A13RAP2.12::GFP in Fig. 2a). Thus, manipulation of the conserved N terminus of RAP2.12 seems to affect the oxygen-dependent subcellular localization of the transcription factor and, moreover, stabilizes the protein under aerobic conditions.

As RAP2.12 has no hydrophobic domains that could explain its localization at the plasma membrane, we searched for interaction partners of the transcription factor. Yeast two-hybrid analyses (Fig. 2b) and bimolecular fluorescence complementation (BiFC) analysis (Fig. 2c) revealed an interaction between RAP2.12 and the membrane-localized acyl-CoA-binding proteins ACBP1 and ACBP2 (ref. 18), as had been shown previously for ACBP2 and RAP2.3 (ref. 19). The interaction between RAP2.12 and ACBP depended on an amino acid sequence between position 123 and 177, which covers the RAYD motif, a sequence already known to mediate protein–protein interactions<sup>20</sup> (Fig. 2d).

The essential role of the N-terminal residues of RAP2.12 is further supported by the conservation of the first amino acids in almost all members of ERF subfamily VII (Supplementary Fig. 8). The specific sequence of their conserved N terminus qualifies ERF-VII proteins as candidate substrates of the N-end rule pathway<sup>21,22</sup> (Fig. 3a).

According to this pathway the terminal Met is removed from the protein by methionine aminopeptidase (MetAP) when the second amino acid of the protein is Cys<sup>23</sup> (Supplementary Fig. 13 and Supplementary Table 5). Terminal Cys is oxidized to cysteine sulphenic acid in an oxygen-dependent manner before arginine transferase (ATE) conjugates an Arg residue to the protein<sup>10,11</sup>. This triggers subsequent ubiquitination by the ligase PROTEOLYSIS 6 (PRT6)<sup>24</sup> and targets the protein to the proteasome for degradation<sup>25</sup>, which can occur in both the cytosol and the nucleus<sup>26</sup>. Transient expression of RAP2.12::GFP in *ate1ate2* or *prt6* knockout plants resulted in accumulation of the transcription factor in the nucleus both during aerobic and hypoxic conditions as well as after re-oxygenation (Fig. 3b), similar to what we observed by deleting the N-terminal sequence (Fig. 2a) or after incubation with the proteasome inhibitor MG132 (Supplementary Fig. 14). Western blot analyses showed that the amount of RAP2.12 increased under hypoxia and decreased again after re-oxygenation in the wild type but not in *ate1ate2* or *prt6* (Fig. 3c). The tolerance to submergence of *ate1ate2* and *prt6* rosette plants was reduced (Supplementary Fig. 15), in line with the negative impact of 35S::HA::RAP2.12 and 35S::A13RAP2.12 on survival (Fig. 1a, b). Lastly, exchanging the N-terminal Cys with Ala (35S::MAG-RAP2.12::GFP) resulted in a GFP signal in the nucleus, similar to what we observed in any of the

**Figure 3 | Oxygen-dependent destabilization of RAP2.12.**

**a**, Graphical representation of the N-end rule branch that leads to oxygen-dependent protein degradation via the 26S proteasome.

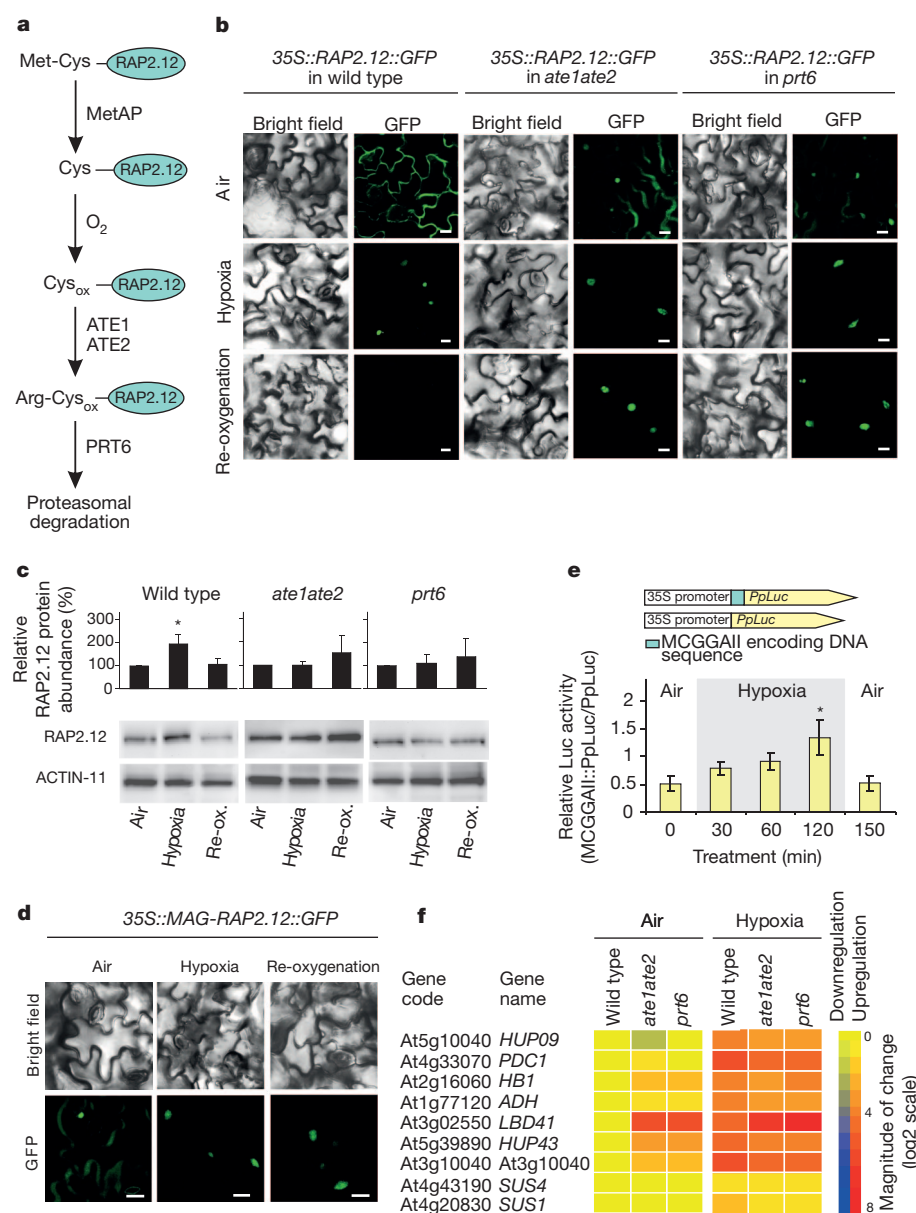
**b**, Subcellular localization of transiently expressed RAP2.12::GFP in leaves ( $n = 9$ ).

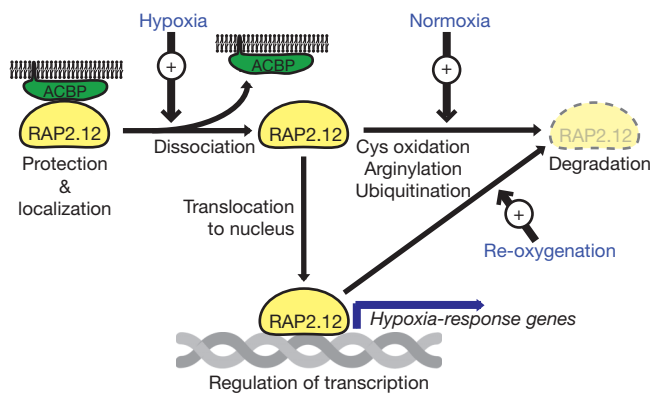
**c**, Western blot and protein quantification of RAP2.12 (means  $\pm$  s.d.,  $*P < 0.05$ , one-way ANOVA,  $n = 3$ ).

Re-ox., re-oxygenation. **d**, Subcellular localization of transiently expressed MAG-RAP2.12::GFP in wild-type leaves ( $n = 9$ ).

**e**, Oxygen-dependent modulation of luciferase after fusion with RAP2.12 N-terminal amino acid residues (means  $\pm$  s.e.,  $*P < 0.05$ , one-way ANOVA,  $n = 6$ ).

**f**, Expression of hypoxia-inducible genes in *ate1ate2* and *prt6* mutants.





**Figure 4 | Model describing the oxygen sensor mechanism in plants.** The transcription factor RAP2.12 is constitutively expressed under aerobic conditions. RAP2.12 protein is always present, bound to ACBP to prevent RAP2.12 from moving into the nucleus under aerobic conditions and to protect it against proteasomal degradation in air. Upon hypoxia, RAP2.12 moves into the nucleus, where it activates anaerobic-gene expression. Upon re-oxygenation, RAP2.12 is rapidly degraded via the N-end rule pathway and proteasome-mediated proteolysis to downregulate the hypoxic response.

other approaches to modify the N-end rule pathway (Fig. 3d). All this indicates that the lifetime of RAP2.12 is controlled by the N-end rule pathway for proteasomal protein degradation.

Next, we investigated whether an oxygen-dependent N-end rule pathway is active in plants and whether it regulates the oxygen-dependent activation of hypoxic gene expression. Fusion of the first conserved N-terminal amino acid residues from RAP2.12 to the Luc reporter protein resulted in an increase of the normalized Luc activity under hypoxic conditions and reduced Luc activity upon re-oxygenation, as predicted by the Cys-oxidation-dependent branch of the N-end rule pathway (Fig. 3e). In addition, constitutive upregulation of hypoxia marker genes was observed under aerobic conditions in plants with reduced ATE and PRT activities (Fig. 3f). This indicates that the oxygen-dependent oxidation of the terminal Cys of RAP2.12 prevents hypoxic gene expression via the destabilization of RAP2.12 in air. Only when the oxygen concentration decreases is Cys oxidation prevented, and the now stably accumulating RAP2.12 can induce the expression of genes involved in the hypoxic response (Fig. 4). Here, we have shown that this oxygen-dependent Cys oxidation is adopted by the ERF-VII factor RAP2.12 and—together with its oxygen-dependent re-localization—triggers the hypoxia-acclimation response in *Arabidopsis*.

## METHODS SUMMARY

Unless specifically indicated in the text, low oxygen (hypoxia) conditions used in this study were always maintained at 1% (v/v) oxygen. Full details of the materials and experimental procedures are provided in Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 19 May; accepted 6 September 2011.

Published online 23 October 2011.

- Webb, J. D., Coleman, M. L. & Pugh, C. W. Hypoxia, hypoxia-inducible factors (HIF), HIF hydroxylases and oxygen sensing. *Cell. Mol. Life Sci.* **66**, 3539–3554 (2009).
- Kelly, D. P. Hypoxic reprogramming. *Nature Genet.* **40**, 132–134 (2008).
- Mustroph, A. *et al.* Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. *Plant Physiol.* **152**, 1484–1500 (2010).

- Bailey-Serres, J. & Voesenek, L. A. C. J. Flooding stress: acclimations and genetic diversity. *Annu. Rev. Plant Biol.* **59**, 313–339 (2008).
- Green, J., Crack, J. C., Thomson, A. J. & LeBrun, N. E. Bacterial sensors of oxygen. *Curr. Opin. Microbiol.* **12**, 145–151 (2009).
- Hou, S. *et al.* Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature* **403**, 540–544 (2000).
- Osborne, T. F. & Espenshade, P. J. Evolutionary conservation and adaptation in the mechanism that regulates SREBP action: what a long, strange tRIP it's been. *Genes Dev.* **23**, 2578–2591 (2009).
- Semenza, G. L. HIF-1, O<sub>2</sub>, and the 3 PHDs: how animal cells signal hypoxia to the nucleus. *Cell* **107**, 1–3 (2001).
- van der Wel, H. *et al.* Requirements for Skp1 processing by cytosolic Prolyl 4(*trans*)-hydroxylase and  $\alpha$ -N-acetylglucosaminyltransferase enzymes involved in O<sub>2</sub> signaling in *Dictyostelium*. *Biochemistry* **50**, 1700–1713 (2011).
- Lee, M. J. *et al.* RGS4 and RGS5 are *in vivo* substrates of the N-end rule pathway. *Proc. Natl Acad. Sci. USA* **102**, 15030–15035 (2005).
- Graciet, E., Mesiti, F. & Wellmer, F. Structure and evolutionary conservation of the plant N-end rule pathway. *Plant J.* **61**, 741–751 (2010).
- Hinz, M. *et al.* *Arabidopsis* RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. *Plant Physiol.* **153**, 757–772 (2010).
- Licausi, F. *et al.* HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana*. *Plant J.* **62**, 302–315 (2010).
- Xu, K. *et al.* *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**, 705–708 (2006).
- Hattori, Y. *et al.* The ethylene response factors *SNORKEL1* and *SNORKEL2* allow rice to adapt to deep water. *Nature* **460**, 1026–1030 (2009).
- Papdi, C. *et al.* Functional identification of *Arabidopsis* stress regulatory genes using the controlled cDNA overexpression system. *Plant Physiol.* **147**, 528–542 (2008).
- Licausi, F. *et al.* Hypoxia responsive gene expression is mediated by various subsets of transcription factors and miRNAs that are determined by the actual oxygen availability. *New Phytol.* **190**, 442–456 (2011).
- Li, H. Y. & Chye, M. L. Membrane localization of *Arabidopsis* acyl-CoA binding protein ACBP2. *Plant Mol. Biol.* **51**, 483–492 (2003).
- Li, H. Y. & Chye, M. L. *Arabidopsis* Acyl-CoA-binding protein ACBP2 interacts with an ethylene-responsive element-binding protein, AtEBP, via its ankyrin repeats. *Plant Mol. Biol.* **54**, 233–243 (2004).
- Okamuro, J. K., Caster, B., Villarreal, R., Van Montagu, M. & Jofuku, K. D. The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **94**, 7076–7081 (1997).
- Kwon, Y. T. *et al.* An essential role of N-terminal arginylation in cardiovascular development. *Science* **297**, 96–99 (2002).
- Graciet, E. & Wellmer, F. The plant N-end rule pathway: structure and functions. *Trends Plant Sci.* **15**, 447–453 (2010).
- Bradshaw, R. A., Brickey, W. W. & Walker, K. W. N-terminal processing: the methionine aminopeptidase and N<sup>acetyl</sup> transferase families. *Trends Biochem. Sci.* **23**, 263–267 (1998).
- Garzón, M. *et al.* *PRT6/At5g02310* encodes an *Arabidopsis* ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the *CER3* locus. *FEBS Lett.* **581**, 3189–3196 (2007).
- Voges, D., Zwickl, P. & Baumeister, W. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* **68**, 1015–1068 (2000).
- Vallon, U. & Kull, U. Localization of proteasomes in plant cells. *Protoplasma* **182**, 15–18 (1994).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We would like to thank H. van Veen and R. Sasidharan (for providing *Rumex* data), E. Maximova, F. Kragler (microscopy), W. Schulze and R. Bock (support and discussion), A. Fernie and R. Pierik (commenting on the manuscript) and S. Parlanti, L. Bartzeko and K. Seehaus (plant cultivation). This work was financially supported by the Max Planck Institute of Molecular Plant Physiology, Scuola Superiore Sant'Anna, and the Deutsche Forschungsgemeinschaft (DFG) (DO 1298/2-1).

**Author Contributions** F.L., M.K., D.A.W. and B.G. performed the experiments. F.M.G. carried out the bioinformatic analysis. F.L., L.A.C.J.V., P.P. and J.T.v.D. designed the experiments. F.L., P.P. and J.T.v.D. wrote the manuscript. All the authors discussed and commented on the content of the paper.

**Author Information** The raw data files of the microarray experiments have been deposited in the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>; accession number: GSE29187). The gene sequences for the *Rumex* spp. used in this work have been deposited at NCBI (*RaERF1*: JF968115; *RaERF2*: JF968116; *RpERF1*: JF968117; *RpERF2*: JF968118; and *RpERF3*: JF968119). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to F.L. (f.licausi@sssup.it) or J.T.v.D. (dongen@mpimp-golm.mpg.de).

## METHODS

**Plant materials.** *A. thaliana* Columbia-0 (Col-0) was used as wild-type ecotype, as described in the figure legends. Double *ate1ate2* knockout seeds were provided by E. Graciet, *prt6* knockout seeds (line EOL4) were obtained from the Institute of Agronomic Research, 35S::GFP seeds were provided by M. Kawai-Yamada.

**Growth conditions and phenotypic evaluation.** Seeds were sown in moist soil, stratified at 4 °C in the dark for 48 h and germinated at 22 °C day/18 °C night with a photoperiod of 8 h light and 16 h darkness. For all experiments 5-week-old plants were used. Low oxygen (1% (v/v) oxygen in air) treatments were performed as described previously<sup>17</sup>. Flooding tolerance was assayed using three independent transgenic lines. Plants were submerged with deionized water in 15-cm-high plastic boxes and kept in the dark. Leaves were at 5 cm under the water surface. After 84 h, the water was removed from the boxes and photoperiodic conditions (8 h/16 h, light/dark) were restored. Tolerance assays were repeated four times by using 10–20 plants per genotype each time. *Rumex spp.* cultivation and submergence treatment were performed as described previously<sup>27</sup>.

**Cloning of the various constructs.** Coding sequences (CDSs) were amplified from a cDNA template using Phusion High Fidelity DNA-polymerase (New England Biolabs). An artificial microRNA (amiRNA) against RAP2.12 was generated by overlapping PCR using the pRS300 vector as backbone. All open reading frames were cloned into pENTR/D-TOPO (Invitrogen). The resulting entry vectors were recombined into destination vectors using the LR reaction mix II (Invitrogen) to obtain the expression vectors. A complete list of all destination vectors and primers used is provided in Supplementary Tables 6 and 7, respectively.

**Plant transformation.** Stable transgenic plants were obtained using the floral dip method<sup>28</sup>. T0 seeds were screened for kanamycin or phosphinotricin resistance and single-insertion lines were identified as described previously<sup>13</sup>. Transient leaf transformations using 3-week-old plants were performed as described previously<sup>29</sup>. All transient expression assays were repeated at least three times using independently grown plants. Each time the experiment was repeated, we transformed leaves from three independent plants. So, at least nine independent transformations from at least three different plant cultures were analysed.

**qRT-PCR.** RNA extraction, removal of genomic DNA, cDNA synthesis and qRT-PCR analyses were performed as described previously<sup>13</sup>. For 35S::RAP2.12, 35S::HA::RAP2.12, amiRAP2.2-12 and 35S::A13RAP2.12 three independent transgenic lines were used and the average expression value was calculated. For all the other genotypes, three independent biological replicates were used.

**Microarrays.** Three independent RAP2.12 overexpressors or RAP2.2-RAP2.12 silenced lines were grown in soil for 5 weeks and then subjected to a treatment with 1% oxygen in the dark for 90 min. Total RNA from whole rosettes was extracted as described for the qRT-PCR analyses. Hybridization and scanning procedures were performed by NASC (<http://arabidopsis.info/>). Microarray analysis and data quality control were performed as described previously<sup>13</sup> using Robin<sup>30</sup>. Normalization of the raw data and an estimation of signal intensities were carried out using the Genechip Robust Multiarray Average (GC-RMA) methodology<sup>31</sup>. Differential gene expression analysis was carried out using limma<sup>32</sup>, with a Benjamini-Hochberg *P*-value correction<sup>33</sup>. Microarray data sets were deposited in a public repository with open access (accession number GSE29187; <http://www.ncbi.nlm.nih.gov/geo/>).

**Confocal imaging.** For GFP and YFP imaging, leaves from independent stable or transiently transformed 4-week-old plants were analysed with a Leica DM6000B/SP5 confocal microscope (Leica Microsystems).

**Reporter transactivation assay.** *Arabidopsis* mesophyll protoplasts were used to identify the region responsible for the trans-activation activity of RAP2.12. The DNA-binding domain from *Saccharomyces cerevisiae* was fused at the N terminus of RAP2.12 and its deletion variants. The UAS fused to a minimal 35S promoter was inserted into pGreenII-800LUC to generate a reporter and normalization vector. Non-recombined pBDB-GW vector was used as a negative control. Protoplasts were prepared according to a previously described method<sup>34</sup> and transfected using 5 µg plasmid DNA each. A dual luciferase reporter assay was performed as described previously<sup>16</sup>.

**Protein stability assay using the Luc reporter system.** Leaves of *A. thaliana* Col-0 were transformed with either a 35S::PpLuc or a 35S::MCGGAIL::PpLuc

constructs (both containing also a 35S::RrLuc cassette for normalization purposes). Normalized luciferase activity (PpLuc/RrLuc) was measured as described previously<sup>17</sup> and Luc protein stability was evaluated as the ratio between MCGGAIL::PpLuc and PpLuc transfected leaves. The experiment was repeated three times using five independent replicates in each repetition.

**Yeast two-hybrid assay.** The ProQuest™ Two-hybrid System (Invitrogen) was used. PExpTM32/Krev1 and pEXPTM22/RaGDS-wt were used as positive controls, and pDESTTM32 and pDESTTM22 as negative controls. *S. cerevisiae* strain Mav203 was transformed with the different combinations of bait, prey and control vectors (Supplementary Fig. 16). Colonies containing both vectors were selected by plating at 28 °C to select colonies containing an interacting protein partner for 3 days on minimal selective dropout medium lacking Leu and Trp (SC-LW medium). They were subsequently replicated on selective dropout medium (SC-LWH+3AT medium) lacking Leu, Trp, His and supplemented with 10 mM 3-aminotriazole (3AT). The strength of the interaction was further verified by β-galactosidase staining (LacZ) following the manufacturer's instructions.

SC-LW, control medium without Leu and Trp; SC-LWH+3AT, selective medium without Leu, Trp, His and with 3AT.

**BiFC.** *In planta* protein interactions were investigated with bimolecular fluorescence complementation in an *Arabidopsis* transient expression system as described previously<sup>35</sup>.

**SDS-PAGE and western blotting.** Protein samples from total tissue extracts were separated by SDS-PAGE on 10% acrylamide midgels (Biorad) and then transferred onto a polyvinylidene difluoride membrane (BioRad). Incubations with the antiserum and the secondary antibody conjugated to horseradish peroxidase (Agrisera) were performed following the method recommended for the ECL Plus western blotting detection system (GE Healthcare).

Polyclonal anti-RAP2.12 antibodies were affinity purified at Genscript laboratories after being raised in rabbits against a RAP2.12/RAP2.2 specific synthetic peptide (NLKGSKKSSKNRSN). Lyophilized antibody was re-suspended to an approximate concentration of 1 µg ml<sup>-1</sup>. A monoclonal antibody against *Arabidopsis* ACTIN-11 (Agrisera, AS10 702) was used to confirm equal loading and transfer. Densitometric analysis of the protein signals on the western blots was performed with the software package UVP VisionWorks LS (Ultra-Violet Products). Normalization was carried out using the ACTIN-11 signal and setting to 100 the relative protein signal value for each of the 'air' controls.

**Statistical analyses.** Significant variations between genotypes or treatments were evaluated statistically by Sigmaplot using either a *t*-test, one-way or two-way ANOVA where appropriate. Mean values that were significantly different (*P* < 0.05) from the control or wild-type treatment are marked with an asterisk. The statistical evaluation of the microarray experiments is described earlier.

27. Pierik, R., de Wit, M. & Voesenek, L. A. C. J. Growth-mediated stress escape: convergence of signal transduction pathways activated upon exposure to two different environmental stresses. *New Phytol.* **189**, 122–134 (2011).
28. Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W. & Chua, N. H. Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols* **1**, 641–646 (2006).
29. Lee, M. W. & Yang, Y. Transient expression assay by agroinfiltration of leaves. *Methods Mol. Biol.* **323**, 225–229 (2006).
30. Lohse, M. et al. Robin: An intuitive wizard application for R-based expression microarray quality assessment and analysis. *Plant Physiol.* **153**, 642–651 (2010).
31. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article 3 (2004).
32. Reiner, A., Yekutieli, D. & Benjamini, Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* **19**, 368–375 (2003).
33. Wu, Z., Irizarry, R. A., Gentleman, R., Martinez-Murillo, F. & Spencer, F. A model-based background adjustment for oligonucleotide expression arrays. *J. Am. Stat. Assoc.* **99**, 909–917 (2004).
34. Yoo, S. D., Cho, Y. H. & Sheen, J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols* **2**, 1565–1572 (2007).
35. Gehl, C., Waadt, R., Kudla, J., Mendel, R.-R. & Hänsch, R. New GATEWAY vectors for high throughput analyses of protein–protein interactions by bimolecular fluorescence complementation. *Mol. Plant* **2**, 1051–1058 (2009).

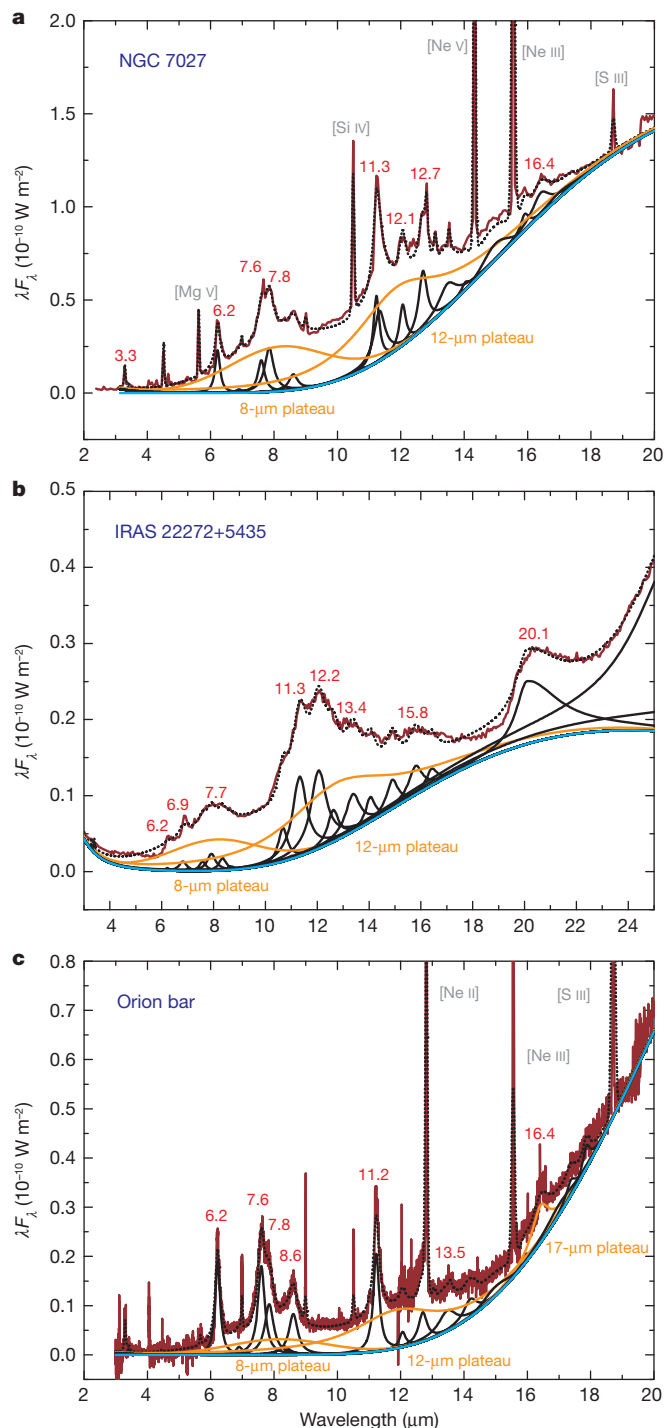
# Mixed aromatic–aliphatic organic nanoparticles as carriers of unidentified infrared emission features

Sun Kwok<sup>1</sup> & Yong Zhang<sup>1</sup>

Unidentified infrared emission bands at wavelengths of 3–20 micrometres are widely observed in a range of environments in our Galaxy and in others<sup>1</sup>. Some features have been identified as the stretching and bending modes of aromatic compounds<sup>2,3</sup>, and are commonly attributed to polycyclic aromatic hydrocarbon molecules<sup>4,5</sup>. The central argument supporting this attribution is that single-photon excitation of the molecule can account for the unidentified infrared emission features observed in ‘cirrus’ clouds in the diffuse interstellar medium<sup>6</sup>. Of the more than 160 molecules identified in the circumstellar and interstellar environments, however, not one is a polycyclic aromatic hydrocarbon molecule. The detections of discrete and broad aliphatic spectral features suggest that the carrier of the unidentified infrared emission features cannot be a pure aromatic compound. Here we report an analysis of archival spectroscopic observations and demonstrate that the data are most consistent with the carriers being amorphous organic solids with a mixed aromatic–aliphatic structure. This structure is similar to that of the organic materials found in meteorites, as would be expected if the Solar System had inherited these organic materials from interstellar sources.

For the past 20 years, polycyclic aromatic hydrocarbon (PAH) molecules have commonly been considered the carriers of unidentified infrared emission (UIE) features. This hypothesis assumes that the UIE features are the result of infrared fluorescence from small (~50-carbon-atom) gas-phase PAH molecules being pumped by far-ultraviolet photons<sup>7</sup>. In spite of its popularity, the PAH hypothesis does not provide

**Figure 1 | Mixed aromatic and aliphatic features in the infrared spectra of circumstellar and interstellar nebulae.** **a–c**, Spectral decompositions of the UIE features of the planetary nebula NGC 7027 (**a**), the proto-planetary nebula IRAS 22272+5435 (**b**) and the Orion bar photodissociation region (**c**), showing a mix of aromatic, aliphatic and continuum features. The observed flux at wavelength  $\lambda$ ,  $F_\lambda$ , is proportional to the emission intensity at that wavelength. A series of discrete features (black lines) and plateau features (orange lines; in **c**, the 17- $\mu\text{m}$  plateau represents the 15–20- $\mu\text{m}$  range) superposed on a continuum (blue line) have been fitted to the observed data. The UIE and plateau features (wavelengths in micrometres), as well as some of the atomic lines, are marked. The observed spectra are shown as solid red lines and the fitted spectra are shown as dotted black lines. The origin of the 20.1- $\mu\text{m}$  feature in the IRAS 22272+5435 spectrum is currently unidentified. The spectral data for NGC 7027, IRAS 22272+5435 and the Orion bar are retrieved from the Infrared Space Observatory archive. For the spectral decomposition, we used the IDL package PAHFIT originally developed to fit the Spitzer Space Telescope Infrared Spectrograph spectra of nearby galaxies. The model spectra take into account the contributions from the stellar continuum, the thermal dust continuum,  $\text{H}_2$  emission, atomic emission lines, the UIE features (both aromatic and aliphatic) and the plateau emission features. The optimal fitting to the observed spectra is achieved through the Levenberg–Marquardt least-squares algorithm. A modified blackbody model for the emission intensity at wavelength  $\lambda$ ,  $I_\lambda \propto \lambda^{-\alpha} B_\lambda(T)$ , where  $B_\lambda(T)$  is the blackbody function with a temperature  $T$ , is used to fit the continuum. The aromatic, aliphatic and plateau features are fitted with assumed Drude profiles  $I_\lambda \propto \gamma^2 [(\lambda/\lambda_0 - \lambda_0/\lambda)^2 + \gamma^2]^{-1}$ , where  $\lambda_0$  is the central wavelength and  $\gamma$  is the fractional full-width at half-maximum of each feature.



<sup>1</sup>Department of Physics, Faculty of Science, The University of Hong Kong, Pokfulam Road, Hong Kong, China.

a good explanation for the observed spectral behaviour. PAH molecules are fused ring molecules made up of carbon and hydrogen, and their vibrational bands are sharp and the peak wavelengths well defined. To fit the broad profiles of the UIE features seen in astronomical spectra, it is necessary to use a complex mixture of PAHs of different sizes, structures and charge states, and to utilize empirical feature profiles<sup>8,9</sup>. Because PAH molecules require ultraviolet photons to excite them, they cannot explain the presence of UIE features in reflection nebulae<sup>10</sup> and proto-planetary nebulae<sup>11</sup> where the central stars are cool and there is no ultraviolet background radiation. To account for these facts, the PAH model has to be revised to include large clusters and other ionization states.

The central argument for the PAH hypothesis is that single-photon excitation of PAH molecules can account for the 12- $\mu\text{m}$  excess emission observed in cirrus clouds in the diffuse interstellar medium by the Infrared Astronomical Satellite (IRAS). However, the UIE-band flux ratios in the diffuse H II regions in the Carina nebula are nearly constant over a range of three orders of magnitude in background radiation<sup>12</sup>. The shapes and peak wavelengths of the UIE features are independent of the temperature of the central stars providing the excitation<sup>10</sup>. Furthermore, PAH molecules have strong and narrow absorption features in the ultraviolet, but these are not observed in interstellar extinction curves<sup>13</sup>. However, the 3.4- $\mu\text{m}$  aliphatic carbon-hydrogen stretching mode is commonly observed in absorption in the diffuse interstellar medium<sup>14</sup>. Although their rotational and vibrational frequencies are well known, not a single PAH molecule has yet been identified in space<sup>15</sup>.

Other arguments have been made to support the PAH hypothesis: the asymmetric profiles of the UIE features can be explained by anharmonicity associated with molecular emission, and the observed feature-to-continuum ratio is high and therefore implies that the carrier is a molecule<sup>7</sup>. Laboratory spectra of mixed aromatic and aliphatic solid materials have asymmetric profiles<sup>16</sup>, which can more

naturally explain the observations. Although the observed feature-to-continuum ratio is high in the diffuse interstellar medium, it is not high in the spectra of planetary and proto-planetary nebulae. Even in the diffuse interstellar medium, the strength of the UIE features are strongly correlated with the dust continuum, suggesting a possible physical relationship between the two components<sup>17</sup>.

The basic premise of the PAH hypothesis does not concern the chemical composition of the carrier so much as its size. As long as the carrier is a nanoparticle that can undergo transient heating, it will satisfy the excitation requirement. Laboratory experiments have yielded carbon nanoparticles with structures of  $sp^2$  rings connected by networks of aliphatic chains<sup>18,19</sup>, as well as fullerene fragments linked by aliphatic groups<sup>20</sup>. These nanoparticles are likely to be constituents in circumstellar and interstellar environments. Furthermore, it has been proposed that the possible sudden release of chemical energy as a source of transient heating of small grains will allow much larger particles to radiate in the near-infrared, further weakening the PAH hypothesis<sup>21</sup>.

An alternative explanation to the UIE bands is that they are emitted by complex organic solids with disorganized structures. These solids intrinsically have broad emission profiles, and the features often sit on even broader emission plateaux several micrometres in width. It has been argued for some time that the observed spectral properties of UIE bands resemble those of coal and kerogen<sup>22,23</sup>. Coal and kerogen are amorphous organic solids with a mixed  $sp^2$ - $sp^3$  composition with randomly oriented aromatic ring units linked by long, aliphatic chains. Their mixed  $sp^2$ - $sp^3$  chemistry gives rise to the discrete aromatic and aliphatic emission features and the broad plateau features<sup>16</sup>.

To provide a quantitative comparison between the two models, we have performed spectral decomposition of several sources with strong UIE features. Figure 1 shows a fit to the infrared spectra of the planetary nebula NGC 7027, the proto-planetary nebula IRAS 22272+5435 and

**Table 1 | Strengths of the UIE discrete and plateau features**

Discrete features* (%)										
	Aromatic					Aliphatic		Unknown		
$\lambda$ ( $\mu\text{m}$ )	3.3	6.2	7.7	8.6	11.3	3.4	6.9	15.8	16.4	18.9
NGC 7027	0.32	1.2	2.8	0.58	3.1	0.07	0.11	0.31	0.84	0.0
IRAS 22272+5435†	0.08	0.05	0.30	0.11	3.76	0.15	0.43	0.94	0.41	—
Orion bar	0.67	3.8	7.0	2.2	2.6	0.13	0.37	0.24	2.0	0.0
V2361 Cygni	0.27‡	0.87§	0.60	0.03	—	—	0.25	—	—	—
V2362 Cygni	—	5.2	2.4	0.8	—	—	3.2	—	—	—
Plateau features (%)										
	Aromatic					Aliphatic		Unknown		
$\lambda$ ( $\mu\text{m}$ )						8	12	17		
NGC 7027						18.8	17.2	0.33		
IRAS 22272+5435						12.5	18.6	—		
Orion bar						9.3	15.1	0.78		
V2361 Cygni						11.1	1.3	—		
V2362 Cygni						17.1	1.2	1.1		
Continuum						Total flux (3-20 $\mu\text{m}$ ) ( $\text{W m}^{-2}$ )				
	Percentage of total flux		Temperature (K)		$\alpha$					
NGC 7027	50.7		100		2	8.9 (−11)				
IRAS 22272+5435	40.8		100		2	2.5 (−11)				
Orion bar	47.3		70		2	2.4 (−11)				
V2361 Cygni	84.8		350		0.2	1.8 (−13)				
V2362 Cygni	66.8		365		0.5	2.1 (−13)				

The UIE phenomenon is complex. In addition to the commonly observed 3.3-, 6.2-, 7.7-, 8.6- and 11.3- $\mu\text{m}$  aromatic features, there are also aliphatic features at 3.4 and 6.9  $\mu\text{m}$ , arising respectively from symmetric and asymmetric carbon-hydrogen stretching and bending modes of methyl and methylene groups attached to aromatic rings. Features at 15.8, 16.4, 17.4 (not shown), 17.8 (not shown), and 18.9  $\mu\text{m}$  have been found in proto-planetary nebulae<sup>11</sup>, reflection nebulae<sup>10</sup> and galaxies. In addition to the discrete features, broad emission features up to several micrometres in width are also seen. The 8- and 12- $\mu\text{m}$  plateau features and a broad feature covering the 15–20- $\mu\text{m}$  range (represented by 17  $\mu\text{m}$  in the table) have been detected in young stellar objects, compact H II regions and planetary nebulae. The 8- and 12- $\mu\text{m}$  plateaux are broad emission features (full-width at half-maximum, 2–4  $\mu\text{m}$ ) and can be identified as collective in-plane and, respectively, out-of-plane bending modes of a mixture of aliphatic side groups attached to aromatic rings<sup>27</sup>. The 15–20- $\mu\text{m}$  plateau feature is also found to be strong in some proto-planetary nebulae. This table summarizes the relative contributions of these components of the spectra shown in Figs 1 and 3.

\*The total fluxes and percentages refer to the values emitted in the 3–20  $\mu\text{m}$  range for NGC 7027 and Orion bar, 5–20  $\mu\text{m}$  for V2362 Cygni and V2361 Cygni, and 3–25  $\mu\text{m}$  for IRAS 22272+5435.

†For IRAS 22272+5435, there are additional features that contribute to the flux in the 3–25- $\mu\text{m}$  range. Features at 12.2, 13.4 and 20.1  $\mu\text{m}$  contribute 3.83, 1.53 and 2.56%, respectively. Some of the contributions also come from the broad 26- $\mu\text{m}$  feature.

‡This entry refers to the spectral feature at 5.3  $\mu\text{m}$ .

§This entry refers to the spectral feature at 6.3  $\mu\text{m}$ .

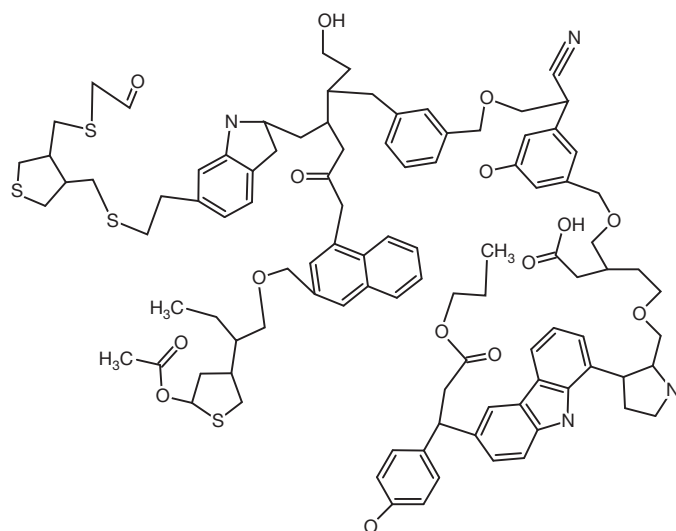
||This entry refers to the spectral feature at 7.2  $\mu\text{m}$ .

the Orion bar, a photodissociation region in the Orion nebula, using a set of discrete UIE features, broad plateau features, and the underlying dust continuum. The breakdown of contributions to the total fluxes from various components is summarized in Table 1. The strongest component is the continuum, which contributes approximately half of the total fluxes emitted in the 3–20- $\mu\text{m}$  region. The next strongest are the plateau features, which account for 36, 31 and 25% of the total fluxes from NGC 7027, IRAS 22272+5435 and the Orion bar, respectively, compared with the totals of 8, 4 and 16% from the aromatic features. The aliphatic branches probably constitute a significant fraction of the material in each of the three sources.

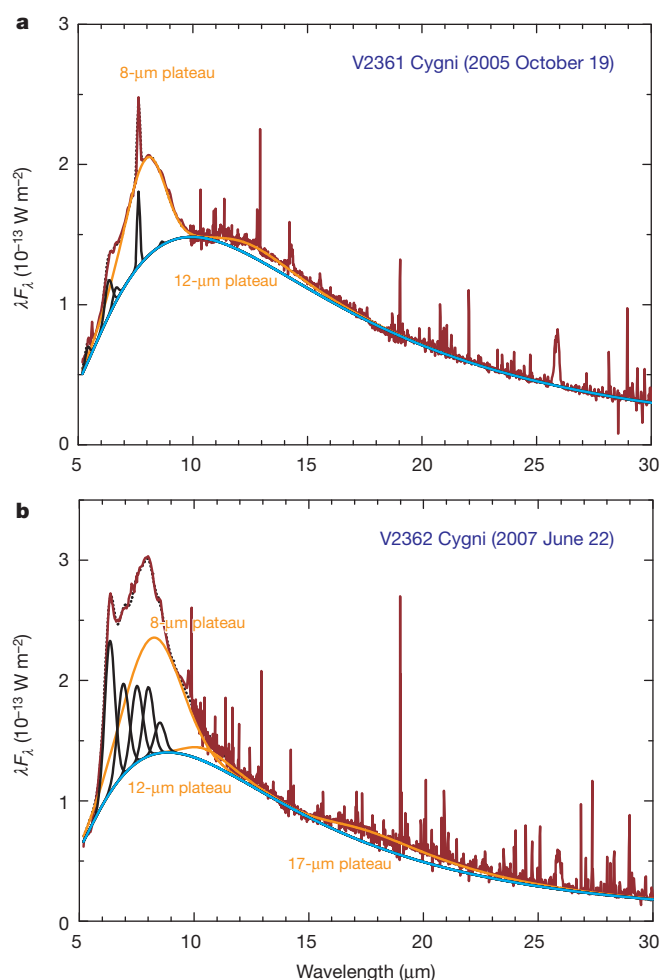
The above fitting results show that the carrier of the UIE features includes a mixture of aromatic and aliphatic components, and is not a pure or predominantly aromatic compound. Because the carrier is formed from a mixture of cosmic gases, it is likely that the compound will include other abundant elements such as oxygen, nitrogen, sulphur and so on, in addition to carbon and hydrogen. These impurities may also have spectral signatures that can be identified by observations at higher spectral resolution. A sketch of the proposed chemical structure is shown in Fig. 2.

The best way to study the origin of the UIE features is to observe them when they are formed. From observations of objects in the late stages of stellar evolution, we know that the UIE features develop in the circumstellar environment within a few hundred years after the termination of the asymptotic giant branch<sup>24</sup>. Spectroscopic observations of novae have shown that the 3.3- and 3.4- $\mu\text{m}$  features appear soon after dust condensation<sup>25</sup>. Theoretically, it is difficult to understand how complex organics can form under such low-density conditions, but novae are observed to change from a pure gas spectrum to a dust-dominated spectrum over the course of days<sup>26</sup>. In Fig. 3, we show a fit to Spitzer spectra of the novae V2362 Cygni and V2361 Cygni. It is expected that a mixture of miscellaneous aliphatic branches will attach to the newly formed ring clusters. The prominence of the plateau features reflects this early stage of organic dust condensation<sup>27</sup>.

We note that the dominant organic content in carbonaceous chondrites is a kerogen-like macromolecular solid referred to as insoluble organic matter. Recent laboratory analysis of the insoluble organic matter in the Murchison meteorite has suggested that it has a chemical structure very similar to that which we propose here<sup>28,29</sup>. The presence of insoluble organic matter in meteorites is evidence that complex



**Figure 2 | Proposed structure of the carrier of UIE features.** The structure is characterized by a highly disorganized arrangement of small units of aromatic rings linked by different kinds of aliphatic chain. Other impurities such as oxygen, nitrogen and sulphur are also commonly present. This structure contains about 100 carbon atoms and a typical nanoparticle may consist of multiple structures similar to this one.



**Figure 3 | Emergence of complex organics after nova outburst.** **a, b,** Fits to the Spitzer Infrared Spectrograph spectra of novae V2361 Cygni (**a**) and V2362 Cygni (**b**) 251 and 446 days after their respective outbursts. In addition to the gas emission line spectrum, both spectra have developed strong dust continua, and the 8- and 12- $\mu\text{m}$  plateau features are clearly present. The continua of V2361 Cygni and V2362 Cygni are fitted by modified blackbody intensities of the respective forms  $\lambda^{-0.2}B_{\lambda}(350\text{ K})$  and  $\lambda^{-0.5}B_{\lambda}(365\text{ K})$  (blue lines). The orange lines are the 8-, 12- and 17- $\mu\text{m}$  plateau features and the solid black lines are discrete features at 5.3, 6.3, 6.9, 7.2 and 8.6  $\mu\text{m}$  for V2361 Cygni and at 6.2, 6.9, 7.6, 7.8 and 8.6  $\mu\text{m}$  for V2362 Cygni. The observed spectra are shown as solid red lines and the fitted spectra are shown as dotted black lines. The presence of the 8- and 12- $\mu\text{m}$  plateau features suggests that the aliphatic component is the first to emerge after dust condensation. Because of the large number of emission lines, the atomic lines are not included in the fitting.

organic solids form in nature with no difficulty. The fact that insoluble organic matter and circumstellar dust have similar chemical structures offers the possibility that Solar System organics may have a stellar connection.

Received 1 April; accepted 30 August 2011.

Published online 26 October 2011.

1. Kwok, S. *Organic Matter in the Universe* (Wiley, 2011).
2. Knacke, R. F. Carbonaceous compounds in interstellar dust. *Nature* **269**, 132–134 (1977).
3. Duley, W. W. & Williams, D. A. The infrared spectrum of interstellar dust: surface functional groups on carbon. *Mon. Not. R. Astron. Soc.* **196**, 269–274 (1981).
4. Allamandola, L. J., Tielens, A. G. G. M. & Barker, J. R. Interstellar polycyclic aromatic hydrocarbons: the infrared emission bands, the excitation/emission mechanism and the astrophysical implications. *Astrophys. J. Suppl. Ser.* **71**, 733–775 (1989).
5. Puget, J. L. & Léger, A. A new component of the interstellar matter: small grains and large aromatic molecules. *Annu. Rev. Astron. Astrophys.* **27**, 161–198 (1989).

6. Sellgren, K. The near-infrared continuum emission of visual reflection nebulae. *Astrophys. J.* **277**, 623–633 (1984).
7. Tielens, A. G. G. M. Interstellar polycyclic aromatic hydrocarbon molecules. *Annu. Rev. Astron. Astrophys.* **46**, 289–337 (2008).
8. Peeters, E. *et al.* The rich 6 to 9  $\mu\text{m}$  spectrum of interstellar PAHs. *Astron. Astrophys.* **390**, 1089–1113 (2002).
9. Draine, B. T. & Li, A. Infrared emission from interstellar dust. IV. The silicate-graphite-PAH model in the post-Spitzer era. *Astrophys. J.* **657**, 810–837 (2007).
10. Uchida, K. I., Sellgren, K., Werner, M. W. & Houdashelt, M. L. Infrared Space Observatory mid-infrared spectra of reflection nebulae. *Astrophys. J.* **530**, 817–833 (2000).
11. Kwok, S., Volk, K. & Hrivnak, B. J. Chemical evolution of carbonaceous materials in the last stages of stellar evolution. *Astron. Astrophys.* **350**, L35–L38 (1999).
12. Onaka, T. Interstellar dust: what do space observations tell us? *Adv. Space Res.* **25**, 2167–2176 (2000).
13. Clayton, G. C. *et al.* The role of polycyclic aromatic hydrocarbons in ultraviolet extinction. I. Probing small molecular polycyclic aromatic hydrocarbons. *Astrophys. J.* **592**, 947–952 (2003).
14. Pendleton, Y. J. & Allamandola, L. J. The organic refractory material in the diffuse interstellar medium: mid-infrared spectroscopic constraints. *Astrophys. J. Suppl. Ser.* **138**, 75–98 (2002).
15. Pilleri, P. *et al.* Search for corannulene ( $\text{C}_{20}\text{H}_{10}$ ) in the Red Rectangle. *Mon. Not. R. Astron. Soc.* **397**, 1053–1060 (2009).
16. Guillois, O., Nenner, I., Papoular, R. & Reynaud, C. Coal models for the infrared emission spectra of proto-planetary nebulae. *Astrophys. J.* **464**, 810–817 (1996).
17. Kahanpää, J., Mattila, K., Lehtinen, K., Leinert, C. & Lemke, D. Unidentified infrared bands in the interstellar medium across the Galaxy. *Astron. Astrophys.* **405**, 999–1012 (2003).
18. Sakata, A., Wada, S., Onaka, T. & Tokunaga, A. T. Infrared spectrum of quenched carbonaceous composite (QCC). II. A new identification of the 7.7 and 8.6 micron unidentified infrared emission bands. *Astrophys. J.* **320**, L63–L67 (1987).
19. Duley, W. W. & Hu, A. Polyyenes and interstellar carbon nanoparticles. *Astrophys. J.* **698**, 808–811 (2009).
20. Jäger, C., Huisken, F., Mutschke, H., Jansa, I. L. & Henning, T. H. Formation of polycyclic aromatic hydrocarbons and carbonaceous solids in gas-phase condensation experiments. *Astrophys. J.* **696**, 706–712 (2009).
21. Duley, W. W. & Williams, D. A. Excitation of the aromatic infrared emission bands: chemical energy in hydrogenated amorphous carbon particles? *Astrophys. J.* **737**, L44 (2011).
22. Papoular, R., Conrad, J., Giuliano, M., Kister, J. & Mille, G. A coal model for the carriers of the unidentified IR bands. *Astron. Astrophys.* **217**, 204–208 (1989).
23. Papoular, R. The use of kerogen data in understanding the properties and evolution of interstellar carbonaceous dust. *Astron. Astrophys.* **378**, 597–607 (2001).
24. Kwok, S. The synthesis of organic and inorganic compounds in evolved stars. *Nature* **430**, 985–991 (2004).
25. Evans, A. *et al.* Infrared spectroscopy of nova Cassiopeiae 1993. IV. A closer look at the dust. *Mon. Not. R. Astron. Soc.* **360**, 1483–1492 (2005).
26. Ney, E. P. & Hatfield, B. F. The isothermal dust condensation of Nova Vulpeculae 1976. *Astrophys. J.* **219**, L111–L115 (1978).
27. Kwok, S., Volk, K. & Bernath, P. On the origin of infrared plateau features in proto-planetary nebulae. *Astrophys. J.* **554**, L87–L90 (2001).
28. Derenne, S. & Robert, F. Model of molecular structure of the insoluble organic matter isolated from Murchison meteorite. *Meteorit. Planet. Sci.* **45**, 1461–1475 (2010).
29. Cody, G. D. *et al.* Establishing a molecular relationship between chondritic and cometary organic solids. *Proc. Natl Acad. Sci. USA* advance online publication, (<http://dx.doi.org/10.1073/pnas.1015913108>) (4 April 2011).
30. Sellgren, K., Uchida, K. I. & Werner, M. W. The 15–20  $\mu\text{m}$  Spitzer spectra of interstellar emission features in NGC 7023. *Astrophys. J.* **659**, 1338–1351 (2007).

**Acknowledgements** We thank A. Tang for technical assistance in the preparation of this manuscript. This work was supported by a grant to S.K. from the Research Grants Council of the Hong Kong Special Administrative Region, China (project no. HKU 7027/11P).

**Author Contributions** S.K. designed the research and wrote the paper. Y.Z. performed data analysis and model fitting.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to S.K. ([sunkwok@hku.hk](mailto:sunkwok@hku.hk)).

# Membrane protein sequestering by ionic protein–lipid interactions

Geert van den Bogaart<sup>1</sup>, Karsten Meyenberg<sup>2</sup>, H. Jelger Risselada<sup>3</sup>, Hayder Amin<sup>1</sup>, Katrin I. Willig<sup>4</sup>, Barbara E. Hubrich<sup>2</sup>, Markus Dier<sup>1</sup>, Stefan W. Hell<sup>4</sup>, Helmut Grubmüller<sup>3</sup>, Ulf Diederichsen<sup>2</sup> & Reinhard Jahn<sup>1</sup>

Neuronal exocytosis is catalysed by the SNAP receptor protein syntaxin-1A<sup>1</sup>, which is clustered in the plasma membrane at sites where synaptic vesicles undergo exocytosis<sup>2,3</sup>. However, how syntaxin-1A is sequestered is unknown. Here we show that syntaxin clustering is mediated by electrostatic interactions with the strongly anionic lipid phosphatidylinositol-4,5-bisphosphate (PIP2). Using super-resolution stimulated-emission depletion microscopy on the plasma membranes of PC12 cells, we found that PIP2 is the dominant inner-leaflet lipid in microdomains about 73 nanometres in size. This high accumulation of PIP2 was required for syntaxin-1A sequestering, as destruction of PIP2 by the phosphatase synaptojanin-1 reduced syntaxin-1A clustering. Furthermore, co-reconstitution of PIP2 and the carboxy-terminal part of syntaxin-1A in artificial giant unilamellar vesicles resulted in segregation of PIP2 and syntaxin-1A into distinct domains even when cholesterol was absent. Our results demonstrate that electrostatic protein–lipid interactions can result in the formation of microdomains independently of cholesterol or lipid phases.

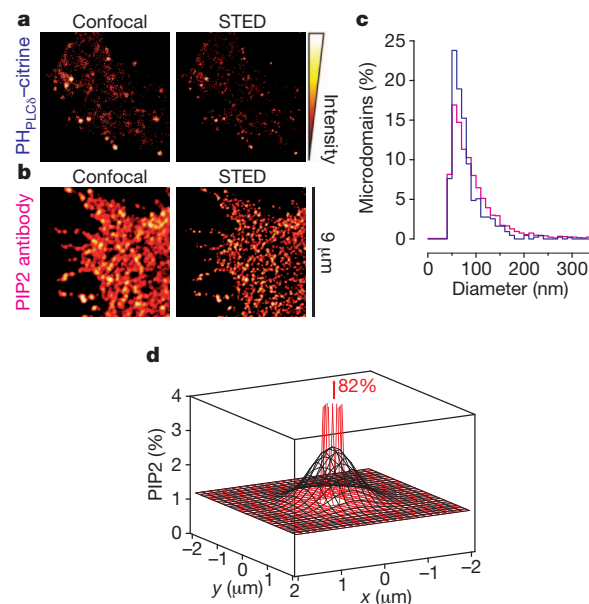
Phosphoinositides are lipids that contain an inositol head group conjugated to one to three phosphate groups. With ~1% of total lipids in the inner leaflet of the plasma membrane<sup>4</sup>, PIP2 is the most abundant phosphoinositide. Earlier studies identified PIP2 as a second messenger in the phospholipase-C signalling pathway. However, the list of cellular functions of PIP2 is rapidly growing, and PIP2 is also involved in membrane targeting, cytoskeletal attachment, endocytosis and exocytosis<sup>4</sup>. PIP2 interacts with many different proteins, through either unstructured basic residue-rich regions or more-structured domains<sup>4,5</sup>.

Neuronal exocytosis requires the presence of PIP2 at the plasma membrane<sup>2,6–8</sup>. The amount of PIP2 at the plasma membrane determines the rate of vesicle priming, the size of the readily releasable pool and the rate of sustained exocytosis in stimulated cells<sup>2,6,8</sup>. This regulation is probably mediated by interactions of PIP2 with proteins involved in docking and fusion such as rabphilin, CAPS, synaptotagmin, SCAMP2 and Mint proteins<sup>6,9</sup>. In docking, PIP2 clusters may act as molecular ‘beacons’ that target synaptic vesicles to the fusion sites. Indeed, PIP2 is locally enriched at the sites of docked vesicles and co-localizes with at least 5–10% of the microdomains of syntaxin-1A<sup>2,3,9</sup> (Supplementary Fig. 1), the membrane-anchored target SNAP receptor of neuronal exocytosis<sup>1</sup>.

The amount of PIP2 at the sites of membrane fusion in PC12 cells has been estimated at 3–6% PIP2 coverage of local cell surface area<sup>9</sup> (Supplementary Fig. 2). In these experiments, membrane sheets were specifically stained for PIP2 with the Pleckstrin homology domain of protein lipase C delta fused to green fluorescent protein<sup>9</sup> or citrine (a yellow fluorescent protein analogue<sup>10</sup>; PH<sub>PLCδ</sub>-citrine; Supplementary Figs 1 and 2), and the fluorescence of the punctuated PIP2 microdomains was quantified. However, this approach underestimates the fraction of PIP2 if the PH<sub>PLCδ</sub> microdomains are smaller than the ~200-nm diffraction-limited resolution of conventional fluorescence

microscopy<sup>9</sup>. To obtain a more accurate estimate, we re-analysed PC12 membrane sheets labelled with PH<sub>PLCδ</sub>-citrine or an antibody raised against PIP2 using super-resolution stimulated-emission depletion (STED) microscopy<sup>11</sup> (Fig. 1a–c). These experiments revealed that the clusters stained for PIP2 are much smaller than anticipated, with an average diameter of only 73 ± 42 nm (s.d.). Although this is an overestimate because it represents the microdomain size convolved with the resolution of the STED microscope (~60 nm), it is in good agreement with the size of the syntaxin-1A microdomains<sup>12</sup>.

Using a microdomain size of 73 nm, we recalculated the surface density of PIP2 (Supplementary Methods). For this calculation, we first estimated the total amount of PIP2 in a microdomain when sampled at the diffraction-limited resolution of our epifluorescence



**Figure 1 | PIP2 is the predominant inner-leaflet lipid in roughly 73-nm-sized microdomains.** **a**, Confocal and corresponding nanoscale-resolution STED image of a PH<sub>PLCδ</sub>-citrine-stained membrane sheet of PC12 cells. Note the increase in resolution. **b**, Same as **a**, but now immunostained with a monoclonal PIP2 antibody and a secondary antibody labelled with Alexa Fluor 488. **c**, Size distribution of microdomains with PH<sub>PLCδ</sub>-citrine (blue;  $n = 433$ , 24 sheets, two independent preparations) and PIP2 antibody (pink;  $n = 2,959$ , 22 sheets, two independent preparations). The average diameter (full-width at half-maximum) was 73 ± 42 nm (s.d.) for microdomains with PH<sub>PLCδ</sub>-citrine and 87 ± 62 nm (s.d.) for microdomains with PIP2 antibody. **d**, Spatial distribution of PIP2. Black: the PIP2 distribution when sampled at too low diffraction-limited resolution (377 nm; Supplementary Fig. 2). Red: approximation of the PIP2 distribution in the ~73-nm microdomains. PIP2 was accumulated over ~82% of the total surface area. See Supplementary Methods for details.

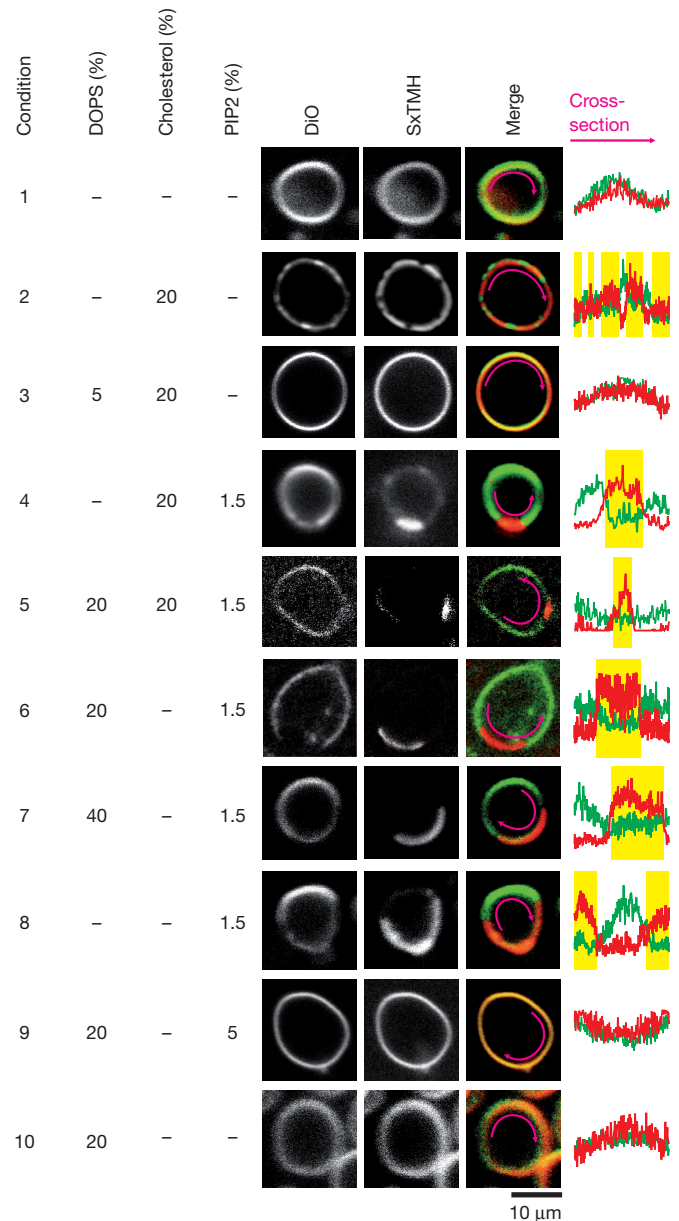
<sup>1</sup>Department of Neurobiology, Max Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany. <sup>2</sup>Institute for Organic and Biomolecular Chemistry, Georg-August-University Göttingen, Tammannstraße 2, 37077 Göttingen, Germany. <sup>3</sup>Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany. <sup>4</sup>Department of Nanobiophotonics, Max Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany.

microscope (Fig. 1d, black mesh, and Supplementary Fig. 2). We then calculated the peak concentration when this amount of PIP2 was concentrated into 73-nm microdomains (Fig. 1d, red mesh). Here we assumed a Gaussian distribution of PIP2 in the microdomains. A peak surface coverage of 82% PIP2 was obtained (Fig. 1d). It must be kept in mind that at these high PIP2 concentrations, molecular crowding might hinder binding of  $\text{PH}_{\text{PLC}\delta}$ -citrine; that relatively small errors in microdomain size and microscope resolution result in a substantial error in the PIP2 coverage; and that  $\text{PH}_{\text{PLC}\delta}$ -citrine and antibody binding may alter PIP2 localization and is only indicative of PIP2 microdomains. Nevertheless, the PIP2 concentrations that we calculate are much higher than any previous estimate, and it seems safe to conclude that PIP2 is the dominant inner-leaflet lipid in the microdomains. The question then arises of by which molecular mechanism such high concentrations of PIP2 are achieved.

PIP2 has a net negative charge of  $-3$  to  $-5$  (ref. 4) and interacts with polybasic stretches of amino acids<sup>4,5,13,14</sup>. Proteins with such stretches can sequester PIP2 in amounts in excess even of monovalent anionic lipids, such as MARCKS, spermine and even pentyllysine<sup>5,14</sup>. Similar to these proteins, syntaxin-1A also possesses a stretch of basic amino acids. These residues are adjacent to the transmembrane domain and are in contact with the head groups of the phospholipids<sup>15,16</sup> (Supplementary Fig. 3a). Indeed, it is well established that this conserved stretch with five positive residues (<sup>260</sup>KARRKK<sup>265</sup>) interacts with PIP2<sup>9,15–17</sup>. Removal of charge weakens this interaction (Supplementary Fig. 3b, c), but syntaxin-1A remains capable of fusing membranes even on removal of all five charges<sup>9,16</sup>. Because PIP2 co-localizes with at least a fraction of syntaxin-1A microdomains<sup>2</sup> (Supplementary Fig. 1), we speculated that their interaction might drive domain formation in a manner similar to that in various soluble lipid-binding proteins<sup>5,14</sup>. Two independent approaches were used to test this hypothesis: reconstitution in giant unilamellar vesicles<sup>18</sup> (GUVs) and hydrolysis of PIP2 in PC12 cells using a membrane-targeted variant of the PIP2 phosphatase synaptojanin-1<sup>7</sup>.

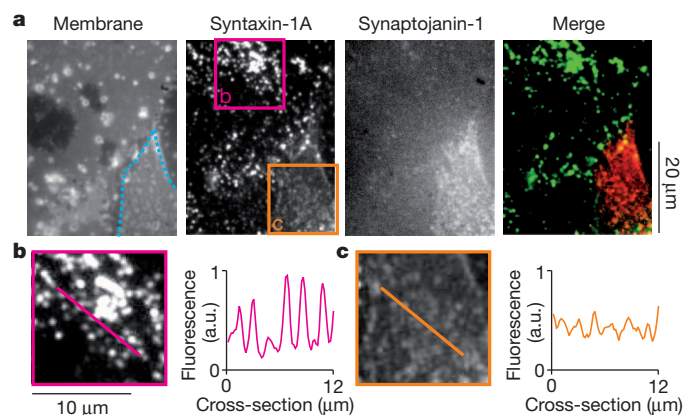
Syntaxin-1A clustered in a non-raft way in neutral cholesterol-phosphatidylcholine membranes<sup>17</sup>. Here cholesterol clusters syntaxin-1A by competing for solvation by phosphatidylcholine. Indeed, a synthetic C-terminal peptide of syntaxin-1A (residues 257–288; 3 mol%; Fig. 2 and Supplementary Fig. 4a) clustered in domains in >50% of GUVs composed of 1,2-dioleoyl-*sn*-glycero-3-phosphatidylcholine (DOPC) with 20 mol% cholesterol. This peptide contained both the polybasic juxtamembrane linker and transmembrane region and was amino-terminally labelled with either rhodamine red or Atto 647N. Analysis of fluorescence showed a  $(1.6 \pm 0.2)$ -fold (s.d.;  $n = 18$ ) enrichment of syntaxin-1A<sub>257–288</sub> in these clusters, but this is an underestimate limited by the optics. Negatively charged PIP2 or 1,2-dioleoyl-*sn*-glycero-3-phosphatidylserine (DOPS) dispersed these clusters<sup>17</sup> (Fig. 2). Thus, although cholesterol competition might explain syntaxin-1A clusters that are not enriched in PIP2 (Supplementary Fig. 1), it cannot explain the high accumulation of PIP2 at the sites of docked vesicles. However, 1.5 mol% (total lipids) PIP2 also clustered syntaxin-1A in 1–10- $\mu\text{m}$  domains in 1–5% of the GUVs (Fig. 2 and Supplementary Fig. 4b, c). These domains did not depend on cholesterol or DOPS. In these domains, PIP2 was  $(1.9 \pm 0.2)$ -fold (s.d.;  $n = 13$ ; Supplementary Fig. 5) enriched and syntaxin-1A<sub>257–288</sub> was  $(5.5 \pm 1.4)$ -fold (s.d.;  $n = 27$ ) enriched, calculated on the basis of fluorescence. Notably, no domains were observed without peptide or when the PIP2 concentration exceeded 5 mol%. Divalent cations can act as bridges between two adjacent lipids and induce aggregation of PIP2 into clusters<sup>19–21</sup>, but even 1 mM  $\text{Ca}^{2+}$  was not sufficient to attract syntaxin-1A away from the microdomains. Domains were present both for synthetic dioleoyl-PIP2 and for PIP2 extracted from pig brain (Supplementary Fig. 4b). Thus, syntaxin-1A can be clustered in the membrane by both cholesterol and PIP2.

These cholesterol- and PIP2-mediated clusters both differ from lipid ‘rafts’. They also differ from each other. First, PIP2 domains are always



**Figure 2 | Confocal microscopy of syntaxin-1A domains in artificial membranes.** Syntaxin-1A<sub>257–288</sub> (SxTMH; red) labelled with Atto 647N was reconstituted in GUVs. The membranes were composed of DOPC with 1.5 mol% of the fluorescent lipid analogue 3,3'-dioctadecyloxycarbocyanine (DiO; green) and the percentages DOPS, cholesterol and PIP2 indicated in the figure. In the absence of anionic phospholipids, 20% cholesterol clustered syntaxin-1A in many small clusters (condition 2), as predicted in ref. 17. Inclusion of >5% anionic DOPS dispersed these clusters (condition 3). PIP2 (1.5%) partitioned SxTMH in 1–10- $\mu\text{m}$ -sized domains regardless of cholesterol or DOPS (conditions 4–8). These clusters were no longer observed with 5% PIP2 (condition 9). The pink arrows show the part of the membrane used for cross-sections. Yellow bars indicate the positions of the domains. More data are presented in Supplementary Figs 4–9.

round and only one or two form per vesicle, whereas cholesterol generally (but not always) induces many small domains (Supplementary Fig. 4). Second, fluorescence recovery after photobleaching showed that syntaxin-1A remained mobile in the PIP2 domains and that syntaxin-1A was essentially immobile in the cholesterol-dependent clusters (Supplementary Fig. 6). Syntaxin-1A thus diffuses in the PIP2 domains and forms large circular domains to minimize boundary energy<sup>21</sup>. Third, 6-dodecanoyl-2-dimethylaminonaphthalene<sup>22</sup> (laurdan) showed that the PIP2 domains were highly hydrated, whereas the



**Figure 3 | Removal of PIP2 reduces syntaxin-1A clustering in PC12 cells.**

**a**, Membrane sheets of PC12 cells stained with 1-(4-trimethylammoniumphenyl)-6-phenyl-1,3,5-hexatriene<sup>12,24</sup> (TMA-DPH). Immunostaining with a monoclonal antibody raised against syntaxin-1A and secondary antibody labelled with DyLight649 showed that endogenous syntaxin-1A clustered in microdomains (region b; pink)<sup>2,3,9,12,24</sup>. Overexpressing the red-fluorescent-protein-tagged and membrane-targeted catalytic region of synaptojanin-1<sup>7</sup> (residues 498–901; cell outlined in blue) reduced this syntaxin-1A clustering 3.7-fold (region c; orange; Supplementary Fig. 13). Synaptojanin-1 is the 5-phosphatase of PIP2 and overexpression of the construct completely removes PIP2 from the membrane (Supplementary Fig. 12). **b**, **c**, Magnifications of regions b (**b**) and c (**c**). Cross-sections along the indicated cuts indicate the clustering. a.u., arbitrary units.

cholesterol domains were much more densely packed (Supplementary Fig. 7). Fourth, phase contrast microscopy showed a thickening of the cholesterol-dependent clusters but not of the PIP2-domains (Supplementary Fig. 8). Thus, even though no saturated lipids are present, the cholesterol-dependent domains have behaviour that resembles the liquid ordered phase. In contrast, the PIP2 domains seem much more disordered and resemble the liquid disordered phase.  $\text{Ca}^{2+}$  demixing of polyanionic amphiphiles showed that electrostatic interactions can indeed lead to liquid-like domains<sup>21</sup>.

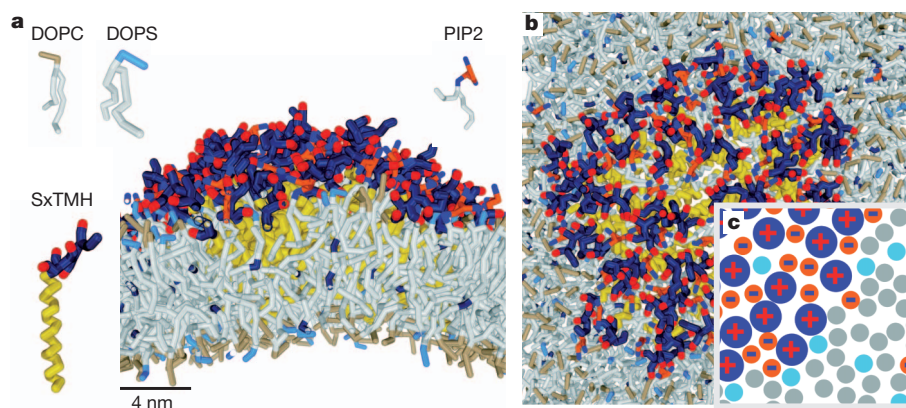
The transmembrane helix of syntaxin-1A has been reported to homodimerize. However, introducing the Met267Ala, Cys271Ala and Ile279Ala mutations that prevent homodimerization of the syntaxin-1A peptides<sup>23</sup> did not prevent cholesterol- or PIP2-mediated clustering (Supplementary Fig. 9). By contrast, no PIP2 domains were observed when two charges (Lys264Ala and Lys265Ala) from the polybasic linker

were removed, but cholesterol-dependent clusters were still observed (Supplementary Fig. 5). Overexpression of the C-terminal part of syntaxin-1A fused to green fluorescent protein<sup>24</sup> in PC12 cells also showed 4–8-fold loss of clustering of the Lys264Ala Lys265Ala mutant (Supplementary Figs 10 and 11). These data show that electrostatic interactions between PIP2 and the juxtamembrane helix of syntaxin-1A are sufficient for domain formation.

We then set out to investigate to what extent PIP2 is required for syntaxin-1A clustering in PC12 cells. For this purpose, we expressed a red-fluorescent-protein-tagged construct containing the phosphatase domain of synaptojanin-1 fused to a CAAX box, resulting in its targeting to the plasma membrane<sup>7</sup>. Synaptojanin-1 is a polyphosphoinositide 5-phosphatase, and the expression of the construct completely removed PIP2 from the plasma membrane<sup>6,7</sup> (Supplementary Fig. 12). Notably, synaptojanin-1 expression reduced 3.7-fold the punctuate distribution of endogenous syntaxin-1A (Fig. 3 and Supplementary Fig. 13). Thus, this provides evidence that PIP2 is indeed required for at least part of syntaxin-1A microdomain formation.

We performed molecular dynamics simulations to gain insight into the precise conformation of the PIP2/syntaxin-1A microdomains. In these coarse-grained simulations, several atoms were represented by one simulation bead<sup>25,26</sup> (Supplementary Fig. 14). This allowed for simulations of relatively large lipid bilayers comprising ~2,500 copies of DOPC and DOPS in a 4:1 molar ratio and 40–64 copies of syntaxin-1A<sub>257–288</sub> and PIP2. In a simulation time of 10  $\mu\text{s}$ , up to ten copies of syntaxin-1A<sub>257–288</sub> clustered with PIP2 into microdomains (Supplementary Fig. 15). Equal amounts of PIP2 and syntaxin-1A were present in the bulk phase of those domains, whereas more PIP2 and DOPS associated transiently to the periphery. We used this information to construct a domain with 64 copies of syntaxin-1A (Fig. 4 and Supplementary Movie 1), which is comparable to the syntaxin-1A content in the microdomains in PC12 cells<sup>12</sup>. These domains were stable over a simulation time of 6  $\mu\text{s}$  and contained <10% residual DOPC or DOPS. We conclude that syntaxin-1A and PIP2 can form dynamic, amorphous networks with PIP2 acting as a ‘charge bridge’ spanning the distance between the various syntaxin-1A molecules (Fig. 4c).

Our findings show that electrostatic interactions between the membrane lipid PIP2 and the SNAP receptor syntaxin-1A suffice to induce membrane sequestering and microdomain formation without the need for high local PIP2 production or a (complex) ‘molecular fence’ restricting PIP2 and protein diffusion<sup>27</sup>. This does not exclude an additional role for protein–protein interactions between either



**Figure 4 | Simulations of the dynamic and amorphous PIP2/syntaxin-1A microdomains.** **a**, **b**, Side view (**a**) and top view (**b**) of a coarse-grained molecular dynamics simulation. Sixty-four copies of syntaxin-1A<sub>257–288</sub> (SxTMH) and 64 copies of PIP2 were incorporated in a bilayer composed of DOPC and DOPS in a 4:1 molar ratio. PIP2 was present only in the membrane leaflet facing the N-terminus of syntaxin-1A<sub>257–288</sub>. Simulations were

performed with 150 mM NaCl. See Supplementary Methods for details. White, lipid alkyl chain; cyan, DOPS head group; grey, DOPC head group; yellow, syntaxin-1A<sub>257–288</sub> transmembrane region (residues 266–288); blue–red, polybasic linker region (residues 257–265; charges in red); orange–blue, anionic PIP2 head group (charges in blue). The domains were stable over a simulation time of 6  $\mu\text{s}$  (Supplementary Movie 1). **c**, Simplified scheme of the cluster.

transmembrane helices or soluble domains. In fact, these seem essential for segregation of proteins with similar structures and sizes, such as syntaxin-1A and syntaxin-4<sup>12,24</sup> (both have polybasic regions and cluster separately). The mutual enrichment of syntaxin-1A and PIP2 at the fusion sites by electrostatic interactions has clear advantages. First, accumulation of syntaxin-1A may facilitate SNAP receptor interactions and thereby increase the membrane fusion efficiency<sup>3,28</sup>. Second, the lipid environment modulates the energetic requirements for fusion<sup>13,16</sup>. Third, both PIP2 and syntaxin-1A function as molecular docking sites and facilitate assembly of the complete fusion machinery<sup>1,2,6–9</sup>. Our findings that electrostatic protein–lipid interactions are sufficient for membrane sequestering indicate that such interactions constitute a mechanism for the formation of protein microdomains in the membrane that is clearly distinct from protein partitioning by means of the well-established lipid phases<sup>18,29</sup>.

## METHODS SUMMARY

PH<sub>PLC8</sub>-citrine was expressed in *Escherichia coli* and purified using His-tag affinity purification. PC12 cells were maintained and propagated as described in refs 3, 24. PC12 cells were transfected using Lipofectamine LTX (Invitrogen). Membrane sheets were prepared by rupturing the cells with probe sonication as described in ref. 24, and immunostaining<sup>24</sup> and microscopy<sup>11,24</sup> were performed as described in the corresponding references. The peptides were synthesized by microwave-assisted Fmoc solid-phase synthesis. Peptides were mixed with lipids in organic solvent and GUVs were formed by the drying rehydration procedure. The molecular dynamics simulations were performed with the GROMACS simulation package and the MARTINI coarse-grained model<sup>25,26</sup>. See Supplementary Methods for details.

Received 24 May; accepted 7 September 2011.

Published online 23 October 2011.

- Jahn, R. & Scheller, R. H. SNAREs – engines for membrane fusion. *Nature Rev. Mol. Cell Biol.* **7**, 631–643 (2006).
- Aoyagi, K. *et al.* The activation of exocytotic sites by the formation of phosphatidylinositol 4,5-bisphosphate microdomains at syntaxin clusters. *J. Biol. Chem.* **280**, 17346–17352 (2005).
- Lang, T. *et al.* SNAREs are concentrated in cholesterol-dependent clusters that define docking and fusion sites for exocytosis. *EMBO J.* **20**, 2202–2213 (2001).
- McLaughlin, S., Wang, J., Gambhir, A. & Murray, D. PIP(2) and proteins: interactions, organization, and information flow. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 151–175 (2002).
- McLaughlin, S. & Murray, D. Plasma membrane phosphoinositide organization by protein electrostatics. *Nature* **438**, 605–611 (2005).
- Wen, P. J., Osborne, S. L. & Meunier, F. A. Dynamic control of neuroexocytosis by phosphoinositides in health and disease. *Prog. Lipid Res.* **50**, 52–61 (2011).
- Milosevic, I. *et al.* Plasmalemmal phosphatidylinositol-4,5-bisphosphate level regulates the releasable vesicle pool size in chromaffin cells. *J. Neurosci.* **25**, 2557–2565 (2005).
- Hay, J. C. & Martin, T. F. Phosphatidylinositol transfer protein required for ATP-dependent priming of Ca<sup>2+</sup>-activated secretion. *Nature* **366**, 572–575 (1993).
- James, D. J., Khodthong, C., Kowalchuk, J. A. & Martin, T. F. Phosphatidylinositol 4,5-bisphosphate regulates SNARE-dependent membrane fusion. *J. Cell Biol.* **182**, 355–366 (2008).
- Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A. & Tsien, R. Y. Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* **276**, 29188–29194 (2001).
- Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).
- Sieber, J. J. *et al.* Anatomy and dynamics of a supramolecular membrane protein cluster. *Science* **317**, 1072–1076 (2007).
- Williams, D., Vicogne, J., Zaitseva, I., McLaughlin, S. & Pessin, J. E. Evidence that electrostatic interactions between vesicle-associated membrane protein 2 and acidic phospholipids may modulate the fusion of transport vesicles with the plasma membrane. *Mol. Biol. Cell* **20**, 4910–4919 (2009).
- Denisov, G., Wanaski, S., Luan, P., Glaser, M. & McLaughlin, S. Binding of basic peptides to membranes produces lateral domains enriched in the acidic lipids phosphatidylserine and phosphatidylinositol 4,5-bisphosphate: an electrostatic model and experimental results. *Biophys. J.* **74**, 731–744 (1998).
- Kweon, D. H., Kim, C. S. & Shin, Y. K. The membrane-dipped neuronal SNARE complex: a site-directed spin labeling electron paramagnetic resonance study. *Biochemistry* **41**, 9264–9268 (2002).
- Lam, A. D., Tryoen-Toth, P., Tsai, B., Vitale, N. & Stuenkel, E. L. SNARE-catalyzed fusion events are regulated by syntaxin1A-lipid interactions. *Mol. Biol. Cell* **19**, 485–497 (2008).
- Murray, D. H. & Tamm, L. K. Clustering of syntaxin-1A in model membranes is modulated by phosphatidylinositol 4,5-bisphosphate and cholesterol. *Biochemistry* **48**, 4617–4625 (2009).
- Bacia, K., Schuette, C. G., Kahya, N., Jahn, R. & Schwiile, P. SNAREs prefer liquid-disordered over “raft” (liquid-ordered) domains when reconstituted into giant unilamellar vesicles. *J. Biol. Chem.* **279**, 37951–37955 (2004).
- Carvalho, K., Ramos, L., Roy, C. & Picart, C. Giant unilamellar vesicles containing phosphatidylinositol(4,5)bisphosphate: characterization and functionality. *Biophys. J.* **95**, 4348–4360 (2008).
- Levental, I. *et al.* Calcium-dependent lateral organization in phosphatidylinositol 4,5-bisphosphate (PIP2)- and cholesterol-containing monolayers. *Biochemistry* **48**, 8241–8248 (2009).
- Christian, D. A. *et al.* Spotted vesicles, striped micelles and Janus assemblies induced by ligand binding. *Nature Mater.* **8**, 843–849 (2009).
- Kaiser, H. J. *et al.* Order of lipid phases in model and plasma membranes. *Proc. Natl Acad. Sci. USA* **106**, 16645–16650 (2009).
- Laage, R., Rohde, J., Brosig, B. & Langosch, D. A conserved membrane-spanning amino acid motif drives homomeric and supports heteromeric assembly of presynaptic SNARE proteins. *J. Biol. Chem.* **275**, 17481–17487 (2000).
- Sieber, J. J., Willig, K. I., Heintzmann, R., Hell, S. W. & Lang, T. The SNARE motif is essential for the formation of syntaxin clusters in the plasma membrane. *Biophys. J.* **90**, 2843–2851 (2006).
- Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI forcefield: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
- Yesylevskyy, S., Schafer, L. V., Sengupta, D. & Marrink, S. J. Polarizable water model for the coarse-grained Martini force field. *PLoS Comp. Biol.* **6**, e1000810 (2010).
- Fujiwara, T., Ritchie, K., Murakoshi, H., Jacobson, K. & Kusumi, A. Phospholipids undergo hop diffusion in compartmentalized cell membrane. *J. Cell Biol.* **157**, 1071–1082 (2002).
- van den Bogaart, G. & Jahn, R. Counting the SNAREs needed for membrane fusion. *J. Mol. Cell Biol.* **3**, 204–205 (2011).
- Simons, K. & Ikonen, E. Functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Holt, G. Bunt, F. S. Wouters and C. Eggeling for advice, and V. Haucke and S. Joo for the red-fluorescent-protein-tagged synaptobrevin-1 construct. G.v.d.B. is financed by the Human Frontier Science Program. This work was supported by the US National Institutes of Health (P01 GM072694, to R.J.) and the Deutsche Forschungsgemeinschaft (SFB803, to K.M., H.J.R., U.D., H.G. and R.J.).

**Author Contributions** G.v.d.B. and R.J. designed the experiments and wrote the paper. K.M., B.E.H. and U.D. synthesized the peptides. H.J.R. and H.G. performed the simulations. K.I.W. and S.W.H. performed the STED microscopy. H.A. and M.D. contributed to the protein purification, immunofluorescence and microscopy. G.v.d.B. performed all other experiments. All authors contributed to writing the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to R.J. ([r.jahn@gwdg.de](mailto:r.jahn@gwdg.de)).

# Image-based genome-wide siRNA screen identifies selective autophagy factors

Anthony Orvedahl<sup>1,2\*</sup>, Rhea Sumpter Jr<sup>1,3\*</sup>, Guanghua Xiao<sup>4</sup>, Aylwin Ng<sup>5,6,7</sup>, Zhongju Zou<sup>1,3,8</sup>, Yi Tang<sup>9</sup>, Masahiro Narimatsu<sup>10</sup>, Christopher Gilpin<sup>11</sup>, Qihua Sun<sup>1,3</sup>, Michael Roth<sup>12,13</sup>, Christian V. Forst<sup>4</sup>, Jeffrey L. Wrana<sup>10,14</sup>, Ying E. Zhang<sup>9</sup>, Katherine Luby-Phelps<sup>11</sup>, Ramnik J. Xavier<sup>5,6,7</sup>, Yang Xie<sup>4,13</sup> & Beth Levine<sup>1,2,3,8,13</sup>

Selective autophagy involves the recognition and targeting of specific cargo, such as damaged organelles, misfolded proteins, or invading pathogens for lysosomal destruction<sup>1–4</sup>. Yeast genetic screens have identified proteins required for different forms of selective autophagy, including cytoplasm-to-vacuole targeting, pexophagy and mitophagy, and mammalian genetic screens have identified proteins required for autophagy regulation<sup>5</sup>. However, there have been no systematic approaches to identify molecular determinants of selective autophagy in mammalian cells. Here, to identify mammalian genes required for selective autophagy, we performed a high-content, image-based, genome-wide small interfering RNA screen to detect genes required for the colocalization of Sindbis virus capsid protein with autophagolysosomes. We identified 141 candidate genes required for viral autophagy, which were enriched for cellular pathways related to messenger RNA processing, interferon signalling, vesicle trafficking, cytoskeletal motor function and metabolism. Ninety-six of these genes were also required for Parkin-mediated mitophagy, indicating that common molecular determinants may be involved in autophagic targeting of viral nucleocapsids and autophagic targeting of damaged mitochondria. Murine embryonic fibroblasts lacking one of these gene products, the C2-domain containing protein, SMURF1, are deficient in the autophagosomal targeting of Sindbis and herpes simplex viruses and in the clearance of damaged mitochondria. Moreover, SMURF1-deficient mice accumulate damaged mitochondria in the heart, brain and liver. Thus, our study identifies candidate determinants of selective autophagy, and defines SMURF1 as a newly recognized mediator of both viral autophagy and mitophagy.

To identify novel genes required for selective autophagy, we performed a genome-wide siRNA screen to detect changes in the colocalization of a red-labelled Sindbis virus (SIN) capsid protein with a green fluorescent protein (GFP)-labelled marker of autophagosomes, GFP-LC3 (LC3 is also known as MAP1LC3) (Supplementary Fig. 1a) in SIN-infected HeLa/GFP-LC3 cells (ref. 6). Using correlative light and electron microscopy, we confirmed that colocalized red and green puncta represented autophagic structures (primarily autolysosomes) containing numerous viral nucleocapsids (Fig. 1a). The predominance of viral nucleocapsids concentrated in these structures (relative to within the cytoplasm) is consistent with selective autophagic targeting of viral nucleocapsids (herein referred to as virophagy).

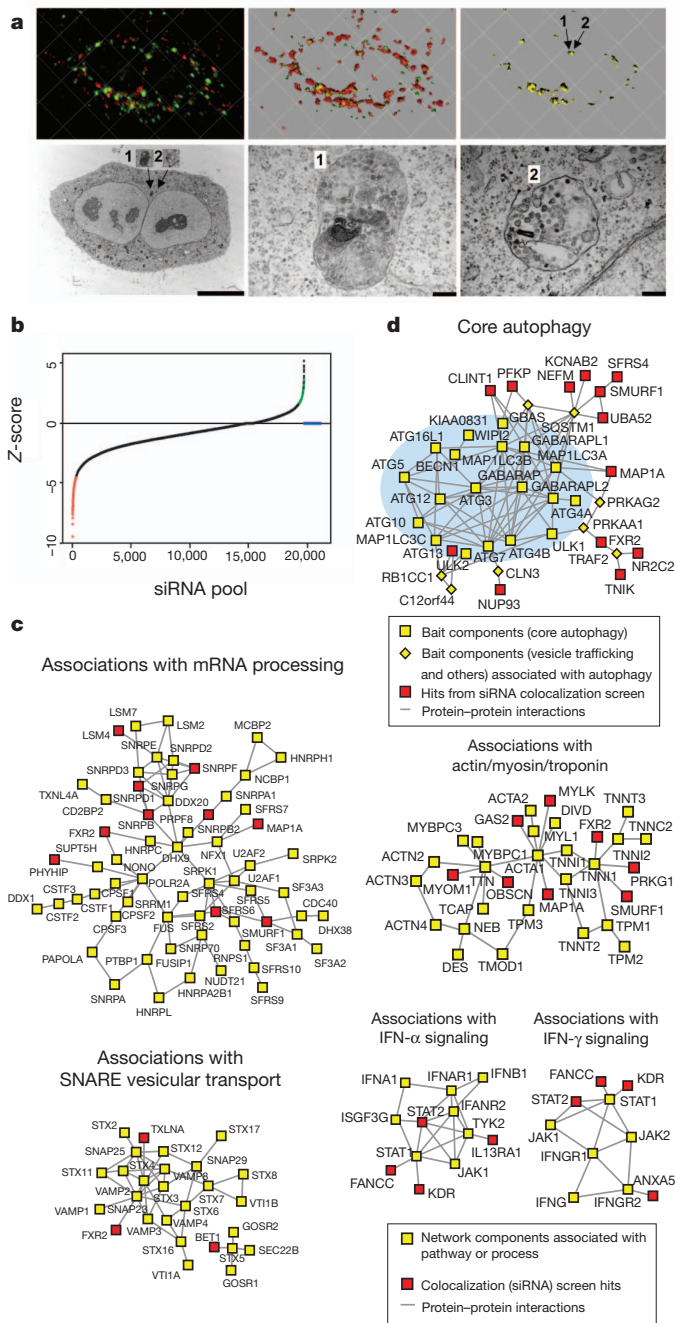
Screening of a human siGenome library containing 21,215 siRNA pools showed that knockdown of 195 and 13 genes resulted in decreased

or increased colocalization, respectively, (Fig. 1b, Supplementary Table 1 and Supplementary Fig. 1b). Genes were re-screened with sets of four individual siRNAs (Supplementary Table 2; see column 'J' of Supplementary Table 3 for siRNA sequences) to confirm our primary screen and rule out potential off-target effects of individual siRNAs; knockdown with two or more siRNAs resulted in decreased colocalization for 141 (72%) genes. (Fig. 2 and Supplementary Figs 1b and 2a). None of these 141 gene knockdowns decreased numbers of green puncta in uninfected cells (data not shown), indicating that these genes function in virophagy, but not in regulation of autophagy. There was no enrichment of siRNAs-containing microRNA seed sequences among these genes ( $P = 0.95$ ) (Supplementary Tables 3 and 4), indicating that bias due to miRNA-like off-target effects was unlikely. There was a low confirmation rate for siRNAs that increased colocalization (2 of 13 genes); therefore, we subsequently focused only on siRNAs that decreased colocalization.

Bioinformatic analyses of the 141 confirmed hits required for SIN capsid/GFP-LC3 colocalization showed enrichment for gene sets associated with biological processes and molecular functions including RNA splicing/processing, protein phosphorylation, transport, calcium-binding and the cytoskeleton (Supplementary Table 5 and Supplementary Fig. 3a). Examination of our hits within a framework of functional cellular pathways revealed strongly enriched network modules associated with RNA processing, interferon (IFN)- $\alpha$  and - $\gamma$  signalling, SNARE vesicular transport, cytoskeletal-associated components, and several metabolic pathways (Supplementary Fig. 3b and Fig. 1c). This is consistent with the function of IFN- $\gamma$  in selective microbial autophagy<sup>1</sup> and the described role of the actin cytoskeleton in selective autophagy in yeast<sup>7</sup> and mammalian<sup>8</sup> cells. The enrichment of SNARE proteins suggests that in addition to a function in autophagosome formation and maturation<sup>9,10</sup>, these proteins may be involved in the trafficking of selective cargo to the autophagosome. Twelve colocalization hits form primary interactions with core autophagy machinery and associated components<sup>11</sup> (Fig. 1d). One colocalization screen hit, clathrin interaction 1 (CLINT) interacts with ATG8 components (GABARAPL1, MAP1LC3A, MAP1LC3B), which are crucial in the recognition of cargo during selective autophagy<sup>2</sup>. Another hit, ATG13, is a member of the core autophagy network, indicating that it may have an as yet undefined function in selective autophagy, in addition to its role in the ULK1 (also known as Atg1) autophagy induction complex<sup>12</sup>. Five colocalization hits, SMURF1, NEFM, KCNAB2, SFRS4 and UBA52, interact with p62 (also known as SQSTM1), a known adaptor in diverse

<sup>1</sup>Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>2</sup>Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>3</sup>Center for Autophagy Research, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>4</sup>Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>5</sup>Center for Computational and Integrative Biology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. <sup>6</sup>Gastrointestinal Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. <sup>7</sup>Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. <sup>8</sup>Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>9</sup>Laboratory of Cellular and Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>10</sup>Center for Systems Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario M5G 1X5, Canada. <sup>11</sup>Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>12</sup>Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>13</sup>Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9113, USA. <sup>14</sup>Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 3E1, Canada.

\*These authors contributed equally to this work.



**Figure 1 | Genome-wide screen to identify cellular factors required for Sindbis virus capsid colocalization with autophagosomes.** **a**, Correlative light and electron microscopy of HeLa/GFP-LC3 cell infected with SIN/mCherry.capsid virus. Top left, deconvolved image of red and green fluorescence channels; middle, three-dimensional surface reconstruction of red and green channels; right, yellow (red + green) colocalization channel. Arrows denote yellow puncta that correspond to '1' and '2' in electron microscopy images below. Bottom left, electron microscopy of identical cell; middle and right, high magnification images of insets '1' and '2'. Scale bars, left, 10  $\mu$ m; middle and right, 200 nm. **b**, Ranked distribution of median Z-scores for each siRNA pool in primary colocalization (virophagy) screen. Red, decreased colocalization; green, increased colocalization; blue, insufficient numbers of green or red puncta per cell or total number of cells per well for analysis. **c**, Maps of protein interactions in enriched network modules (see Supplementary Fig. 3b). **d**, Association of siRNA hits with autophagy network.

forms of selective autophagy<sup>2</sup>, including SIN capsid targeting to autophagosomes<sup>6</sup>.

Selective SIN autophagy (virophagy) promotes the survival of SIN-infected cells<sup>6</sup>. To determine if our identified candidate virophagy

genes have a similar function, we screened our confirmation siRNA library for genes that decreased cell survival after SIN infection. Two or more siRNAs targeting 98 of the genes decreased cell survival after SIN infection (Fig. 2, Supplementary Tables 3 and 6 and Supplementary Figs 1b and 2a); colocalization and cell survival effects of individual siRNAs were significantly correlated (Supplementary Fig. 2b) ( $P = 3.8 \times 10^{-8}$ , Spearman correlation). This is consistent with a pro-survival function of autophagic targeting of SIN capsid in virally infected cells<sup>6</sup>.

To investigate whether the identified candidate virophagy genes also function in other forms of selective autophagy, we performed a secondary screen for autophagy of damaged mitochondria (mitophagy). We used HeLa cells that express an mCherry fusion of Parkin, a cytosolic E3 ubiquitin ligase that translocates to depolarized mitochondria to induce mitophagy after treatment with uncoupling agents (such as CCCP, carbonyl cyanide *m*-chlorophenylhydrazone)<sup>13</sup>. Of the 141 confirmed colocalization hits, 2 or more siRNAs targeting 96 (68%) genes decreased mitophagy (Fig. 2, Supplementary Tables 3 and 7 and Supplementary Figs 1b and 2a, b). Host factors involved in viral autophagy and mitophagy overlapped significantly ( $P = 0.019$ , Spearman correlation). The minority of genes that only scored positive in either the virophagy confirmation or the mitophagy secondary screen may have a role in targeting some, but not other cargoes, for selective autophagy; however, the lack of overlap may also reflect different sensitivities of the two screens. Mitophagy hits consisted of several mitochondria-associated components<sup>14</sup> (NME2, MDH1, NTHL1, PDK1, COX8A, MRPS2, MRPS10, NDUFB9 and BLOC1S) and interactors of mitochondria-associated components (Supplementary Fig. 4).

We focused further on one gene, *SMURF1* (SMAD specific E3 ubiquitin protein ligase 1), encoding a HECT-domain ubiquitin ligase that targets several cytoplasmic proteins for degradation<sup>15</sup>. *SMURF1* was a confirmed hit in all three confirmation or secondary screening assays (see Supplementary Fig. 5 for representative raw data from colocalization confirmation screen), is present in two of the enriched networks (mRNA processing and actin cytoskeleton) (Fig. 1c), and is a predicted interacting partner of the autophagy adaptor, p62 (refs 2, 4, Fig. 1d).

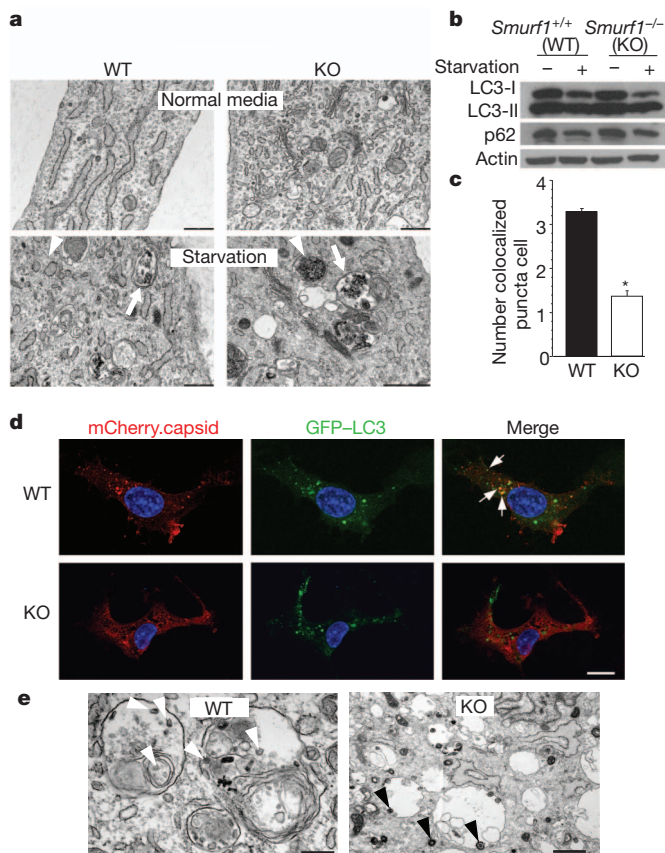
We confirmed that *SMURF1* is not required for general autophagy, but is a bona fide mediator of selective autophagy, including virophagy and mitophagy. siRNA knockdown of *SMURF1* in HeLa cells, unlike knockdown of the essential autophagy protein, ATG7, did not alter general starvation-induced autophagy (Supplementary Fig. 6a). Furthermore, *Smurf1*<sup>-/-</sup> murine embryonic fibroblasts (MEFs) had normal levels of starvation-induced LC3-II (lipidated form of MAP1LC3) conversion, p62 degradation, and ultrastructural evidence of autophagosome and autolysosome accumulation (Fig. 3a, b).

However, a significant decrease in SIN/mCherry.capsid/GFP-LC3 colocalization was observed in SIN-infected *Smurf1*<sup>-/-</sup> MEFs (Fig. 3c, d). Similar to p62 (ref. 6), *SMURF1* and SIN capsid protein co-immunoprecipitate in SIN-infected MEFs and HeLa cells (Supplementary Fig. 7a, b). *SMURF1* is not required for the interaction between p62 and SIN capsid (Supplementary Fig. 7c). The interaction between *SMURF1* and SIN capsid may be relevant for targeting SIN capsid for autophagosomal degradation, as levels of SIN capsid were increased in *Smurf1*<sup>-/-</sup> MEFs and *SMURF1* siRNA-treated HeLa cells. (Supplementary Fig. 7d–f). Increased SIN capsid levels in *Smurf1*-deficient cells cannot be explained by increased capsid production because viral growth was similar in *Smurf1*<sup>-/-</sup> and wild-type MEFs (Supplementary Fig. 7g, h), or by changes in proteasomal degradation because SIN capsid levels were not altered by treatment with the proteasome inhibitor MG132 (Supplementary Fig. 7d), and SIN capsid ubiquitination was not detected (data not shown). SIN-infected *Smurf1*<sup>-/-</sup> MEFs had accelerated cell death (despite similar viral titres) as compared to wild-type controls (Fig. 3g). Thus, *SMURF1* interacts with SIN capsid, *SMURF1* is required for SIN capsid

Gene symbol	C	S	M	Gene symbol	C	S	M	Gene symbol	C	S	M	Gene symbol	C	S	M	Gene symbol	C	S	M	Gene symbol	C	S	M	Gene symbol	C	S	M	Gene symbol	C	S	M
ACIN1	4	4	3	CALCB	0	1	3	EIF2AK1	2	0	1	HIST1H3H	3	2	0	MAP1A	3	2	4	NR3C1	0	0	1	RGS17	1	1	1	STK32A	4	0	2
ACTR1T	2	2	4	CAPS	2	0	2	EIF2S1	0	2	4	HPR	4	3	0	MAP2K1	3	0	4	NTLH1	2	3	4	RIMS3	2	0	4	STOM	2	4	2
ADAMTS1	3	1	3	CCDC36	4	3	1	ENCT	2	0	1	HSF2BP	2	3	0	MAP3K12	3	0	4	NTN4	2	1	1	SATB1	3	1	0	STX10	3	1	1
ADRB2	1	0	2	CD163L1	2	4	3	FABP1	0	4	2	IL13RA1	2	0	1	MBD5	4	4	2	NUP93	4	3	3	SCN1A	1	2	1	SUP3H	3	2	3
AKR1E2	3	1	3	CD93	2	3	3	FAM131B	4	2	3	IMP3	3	1	0	MDH1	4	3	3	OBSCN	3	2	2	SCRN1	1	0	1	TBC1D5	1	1	4
ALKBH5	4	4	1	CDK2AP1	1	1	1	FAM13B	2	2	4	IPPK	3	0	4	MEX3C	2	2	2	P2RX5	2	2	3	SERPINE10	3	4	2	TEAD4	1	3	1
ALPK1	3	3	3	CETN1	2	1	0	FAM176B	1	0	2	ITPKC	1	2	3	MLL3	3	1	1	P2RY4	1	0	0	SFRP4	4	0	3	TMEM203	2	4	2
ANXA5	3	3	2	CHAF1B	1	4	3	FANCC	3	0	3	KCNAB2	4	0	1	MRPS10	4	4	2	PDK1	1	1	4	SFRS4	2	0	1	TMEM39A	0	2	4
ASB2	3	4	4	CHCHD8	2	1	1	FANCF	4	1	2	KCNH3	0	0	1	MRPS2	3	2	3	PDK4	4	2	1	SLC1A3	2	4	3	TMEM39B	3	3	4
ATG13	2	1	2	CHST3	2	2	3	FANCL	2	4	1	KCNK3	2	0	4	MS4A4A	1	0	1	PEX13	2	3	3	SLC1A4	3	0	2	TNKK	2	2	0
ATP1B1	3	4	4	CLDN7	2	2	1	FCGR3B	1	2	4	KCNQ1	2	2	1	MSTN	3	3	3	PEX3	2	4	3	SLC22A3	1	0	3	TREM1	1	4	2
BET1	2	0	0	CLINT1	2	0	1	FGF14	3	4	0	KDR	3	1	0	MTSS1	1	2	1	PFKP	2	1	3	SLC25A19	3	3	0	TRPC5	3	0	0
BLOC1L1	0	0	2	CLVS1	1	1	2	FGF7	2	2	1	KIAA0174	1	3	2	MYH11	0	0	2	PGK2	2	1	4	SLC25A36	3	1	1	TXLNA	2	1	2
BMP2K1	4	2	2	CNOT7	1	0	1	FGFBP1	2	3	3	KIAA0232	0	0	0	MYLK	3	2	3	PHYHIP	3	3	3	SLC35B3	3	3	4	UBA52	2	2	1
BOC	3	1	2	COX6B1	1	1	1	FLJ25363	2	0	2	KRCC1	2	1	4	MYLK3	3	1	0	PI4K2A	2	1	2	SLC35C1	3	4	4	YIPF1	2	3	3
C11orf41	3	3	3	COX8A	0	3	4	FXR2	2	2	1	KRT15	1	1	2	MYOM1	4	3	4	PIK3CA	1	0	4	SLC37A4	3	2	3	ZCCHC17	2	3	3
C14orf104	4	3	2	CPA3	4	2	1	GABRA5	3	3	2	KRT6A	3	3	4	NADSYN1	2	0	1	PLD2	1	1	3	SLC6A1	1	0	4	ZFYVE16	2	1	1
C1orf210	4	3	0	CRINK1	2	3	3	GAS2	2	0	0	KRT73	2	1	3	NAGA	0	0	1	PNPO	2	4	4	SLCO1A2	1	1	2	ZNF189	1	3	2
C1orf223	2	4	2	CSPG5	4	4	2	GDF5	1	2	1	LARP1B	2	1	3	NDUFA4L2	1	1	0	PPY	3	4	2	SMURF1	2	2	2	ZNF593	2	1	4
C2orf12	1	1	2	CXCR7	0	4	1	GMIP	2	2	2	LENG9	3	3	3	NDUFB9	3	2	2	PRKD2	4	0	2	SNRBP	4	3	3	ZNF681	3	2	1
C3orf72	2	1	1	DAAM2	2	0	1	GNE	1	0	0	LMCD1	4	2	4	NEFM	3	0	2	PRKG1	2	0	2	SNRBP2	3	3	0				
C5	3	1	4	DHX38	1	1	0	GPC1	1	0	1	LOC387849	2	3	4	NF2	3	1	1	RBM18	1	3	0	SNRBP1	3	2	4				
C7orf68	2	2	1	DKKL1	4	3	2	GPR81	3	3	2	LOC402210	2	0	0	NLRP14	1	1	1	RCTP1	1	0	0	SNRPF	4	0	2				
C8orf59	3	2	2	DPF3	2	3	4	HAPLN1	1	1	2	LRPAP1	1	0	0	NME2	1	1	4	REEP2	4	2	1	SNTG1	0	2	2				
CA7	2	2	3	DUSP22	0	1	0	HEATR6	3	1	1	LSM4	4	4	3	NR2C2	2	1	4	RFWD3	4	3	2	STAT2	3	4	2				

**Figure 2 | Gene list for viral capsid/autophagosome colocalization (C) confirmation screen and secondary screens for survival of virus-infected cells (S) and Parkin-mediated mitophagy (M).** Shown are the numbers of individual siRNAs from a pool of four targeting each gene that scored positive

in each screen. Red, genes with 2 or more positive siRNAs (confirmed hits); Green, genes with <2 positive siRNAs. See Supplementary Tables 2–8 for further details.



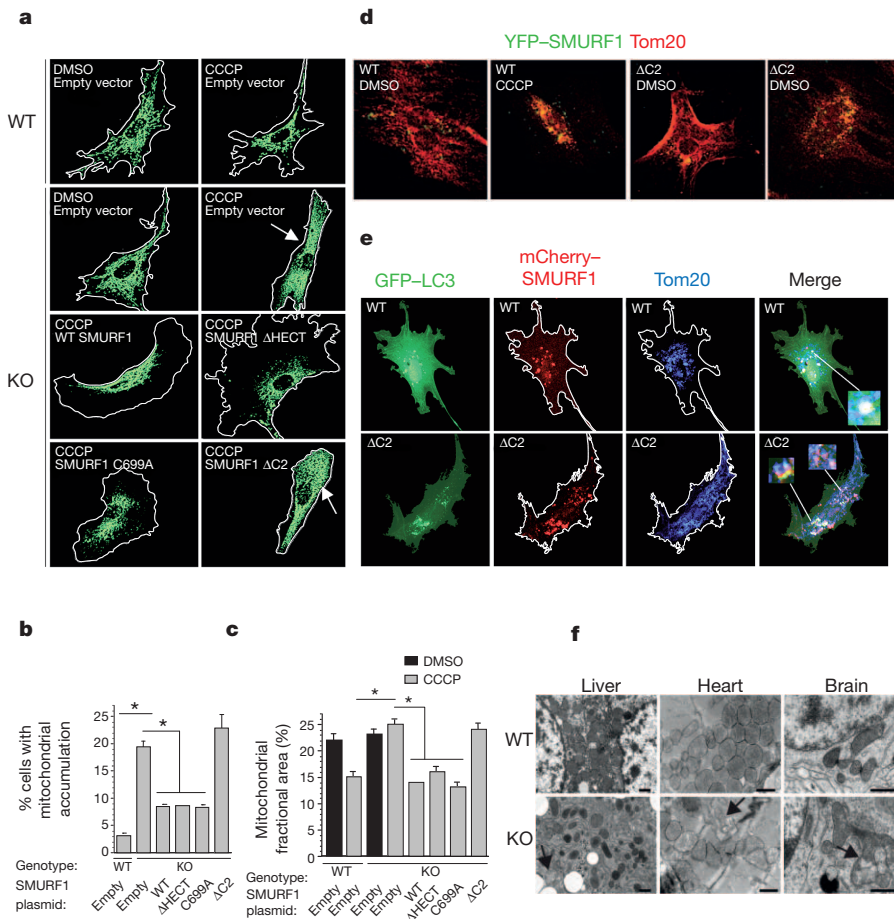
**Figure 3 | SMURF1 functions in virophagy but not in starvation-induced autophagy.** **a**, Electron microscopy analysis of wild-type (WT) and *Smurf1*<sup>-/-</sup> (KO) MEFs grown in normal media or EBSS (starvation) for 4 h. Arrowheads, representative autophagosomes; arrows, representative autolysosomes. Scale bars, 500 nm. **b**, Western blot analyses of LC3-I/II (non-lipidated and lipidated forms of MAP1LC3, respectively) and p62 levels in MEFs of indicated genotype. **c**, Quantification of colocalization of SIN/mCherry.capsid and GFP-LC3 in indicated MEFs 16 h after SIN/mCherry.capsid/GFP-LC3 infection. Data shown represent mean  $\pm$  s.e.m. number of colocalized red and green puncta per cell for 50 cells per well in triplicate samples. \* $P$  < 0.001 against wild-type, Student's  $t$ -test. **d**, Representative confocal microscopy images used for quantification in **c**. Arrows, colocalized red and green puncta. Scale bar, 15  $\mu$ m. **e**, Representative electron microscopy images of indicated MEFs infected with HSV-1 (strain 17termA). White arrowheads, partially degraded viral nucleocapsids inside autolysosomes; black arrowheads, intact viral nucleocapsids inside viral vesicles. Scale bar, 0.5  $\mu$ m. For **a–e**, similar results were obtained in 3–5 independent experiments.

targeting to autophagosomes and degradation through a proteasome-independent pathway, and SMURF1-dependent degradation of SIN capsid promotes cell survival.

To determine whether SMURF1 is required for the autophagic targeting of other viruses, we performed electron microscopy of wild-type and *Smurf1*<sup>-/-</sup> MEFs infected with a mutant strain of herpes simplex virus type 1 harbouring a deletion of ICP34.5, a potent inhibitor of viral autophagy<sup>16–18</sup> (Fig. 3e). As reported<sup>17</sup>, the majority of cytoplasmic HSV-1 virions in wild-type MEFs were inside autolysosomal structures and appeared partially degraded. In contrast, in *Smurf1*<sup>-/-</sup> MEFs, the majority of cytoplasmic HSV-1 virions were inside single-membraned vesicles involved in HSV-1 cytoplasmic egress and had an intact structure. This lack of autophagic targeting of HSV-1 in *Smurf1*<sup>-/-</sup> MEFs was not due to a general defect in autophagy, because HSV-1 infection induced autophagy similarly in *Smurf1*<sup>-/-</sup> and wild-type MEFs (Supplementary Fig. 6b). Thus, SMURF1 is required for the autophagic targeting of both a positive-strand RNA (Sindbis) and a double-stranded DNA (herpes simplex) virus.

Next, we examined the role of SMURF1 in mitophagy. In HeLa cells, all four *SMURF1* siRNAs decreased SMURF1 protein expression (Supplementary Fig. 8a) and inhibited Parkin-mediated CCCP-induced mitophagy as effectively as an siRNA targeting *p62*, a mediator of mitophagy in some previous reports<sup>19,20</sup>, and siRNA targeted against the essential autophagy gene, *ATG7* (Supplementary Fig. 8b, c). The magnitude of each individual siRNA's effect on SMURF1 protein expression knockdown correlated with the magnitude of inhibition of Parkin-mediated autophagy. Therefore, in the mitophagy confirmation screen, two of the *SMURF1* siRNAs were probably false negatives; indeed, the number of Parkin-expressing cells in wells treated with these siRNAs was low (data not shown), precluding meaningful statistical analyses. A similar finding was true in the viral colocalization screen.

We further examined the role of SMURF1 in mitophagy by assessing mitochondrial clearance in CCCP-treated *Smurf1*<sup>-/-</sup> MEFs. Unlike in HeLa cells, Parkin overexpression did not promote mitophagy in MEFs of either genotype (data not shown). However, 25–30% of MEFs treated with 10  $\mu$ M CCCP showed changes in mitochondrial morphology. In wild-type MEFs with damaged mitochondria (swollen or fragmented appearance), partial mitochondrial clearance occurred with compaction of the remaining mitochondria around the nucleus (Fig. 4a). In contrast, in *Smurf1*<sup>-/-</sup> cells with damaged mitochondria, virtually no mitochondrial clearance occurred and there was diffuse accumulation of fragmented mitochondria throughout the cytoplasm (Fig. 4a, arrows). This phenotypic difference was confirmed using two independent methods of quantification, including assessment of the total percentage of CCCP-treated cells that displayed diffuse



**Figure 4 | SMURF1 functions in mitophagy.**

**a**, Representative mitochondrial morphology in *Smurf1*<sup>+/+</sup> (wild-type) and *Smurf1*<sup>-/-</sup> (KO) MEFs transfected with indicated construct and treated with DMSO or 10  $\mu$ M CCCP for 24 h.

**b**, Quantification of percentage of total cells with a diffuse accumulation of abnormal fragmented mitochondria and lack of mitochondrial clearance. Results shown represent combined data from 3–5 experiments per condition with triplicate wells (of at least 100 cells per well) analysed for each condition per experiment. Shown are mean  $\pm$  s.e.m. for average values from each experiment. Similar results were observed in each independent experiment. \**P* < 0.001, Student's *t*-test. **c**, Measurement of mitochondrial fractional area (percentage of total cellular area) in MEFs treated as in **a**. Results shown represent mean  $\pm$  s.e.m. for 50 cells per condition.

**d**, Representative confocal micrographs of KO MEFs transfected with YFP-SMURF1 wild-type or YFP-SMURF1 $\Delta$ C2 ( $\Delta$ C2) and treated for 4 h with DMSO or CCCP. **e**, Representative confocal micrographs of KO MEFs transfected with GFP-LC3 and wild-type mCherry-SMURF1 (WT) or mCherry-SMURF1 $\Delta$ C2 ( $\Delta$ C2) and treated for 4 h with CCCP. Inset, upper right, formation of completed autophagosome around a damaged mitochondrion associated with wild-type SMURF1; insets, lower right, incomplete autophagosomes or absence of LC3 signal around mitochondria associated with SMURF1 $\Delta$ C2. See also Supplementary Figs 10 and 11 for enlarged images. **f**, Representative electron microscopy images of indicated tissues from 10-month-old *Smurf1*<sup>+/+</sup> (WT) or *Smurf1*<sup>-/-</sup> (KO) mice.

Arrows, representative abnormal mitochondria. Scale bars, 1  $\mu$ m. Similar abnormalities were observed throughout entire electron microscopy tissue section for each mouse and in tissue samples from three different mice for each genotype.

accumulation of abnormal mitochondria (Fig. 4b) and the measurement of fractional mitochondrial surface area per cell (Fig. 4c).

To evaluate the mechanism of action of SMURF1, we compared the effects of wild-type and mutant SMURF1 expression plasmids on rescue of selective autophagy in *Smurf1*<sup>-/-</sup> MEFs (Fig. 4a–c). We focused on mitophagy rather than SIN capsid viroplasm because of the resistance of MEFs to SIN infection after plasmid transfection. The defect in mitophagy in *Smurf1*<sup>-/-</sup> MEFs was partially rescued by wild-type SMURF1 transfection. SMURF1 $\Delta$ HECT<sup>21</sup>, lacking the HECT domain that catalyses ubiquitin ligation onto target proteins, or SMURF1(C699A)<sup>21</sup>, a catalytically inactive point mutant, rescued the mitochondrial clearance defect as efficiently as wild-type SMURF1. Thus, in addition to its known role in targeting proteins for proteasomal degradation via ubiquitination, SMURF1 has a ubiquitin ligase activity-independent function in mediating the selective degradation of damaged mitochondria.

In contrast, a SMURF1 mutant lacking the C2 domain, SMURF1 $\Delta$ C2, was completely defective in mitophagy rescue in CCCP-treated *Smurf1*<sup>-/-</sup> MEFs (Fig. 4a–c), despite similar levels of expression as transfected wild-type SMURF1 (Supplementary Fig. 9a). The C2 domain of SMURF1 was not required for SMURF1 co-immunoprecipitation with p62 (Supplementary Fig. 9b), indicating that SMURF1 does not function in selective autophagy by recruiting p62. C2 domains (including those of protein kinase C and SMURF1) bind membrane phospholipids and function in protein targeting to the plasma membrane and/or membrane subcellular compartments<sup>22,23</sup>. This raised the possibility that SMURF1 might function in the targeting of selective autophagy cargo through interaction with the nascent autophagosome membrane.

To investigate this possibility, we examined the subcellular localization of wild-type SMURF1 and SMURF1 $\Delta$ C2 with damaged mitochondria and autophagosomes (Fig. 4d). In *Smurf1*<sup>-/-</sup> MEFs transfected with wild-type, yellow fluorescent protein-conjugated YFP-SMURF1, CCCP treatment induced the colocalization of YFP-SMURF1 with damaged mitochondria. In *Smurf1*<sup>-/-</sup> MEFs transfected with YFP-SMURF1 $\Delta$ C2, increased numbers of fragmented and swollen mitochondria were observed in basal conditions and these increased further upon CCCP treatment. These abnormal mitochondria colocalized with YFP-SMURF1 $\Delta$ C2, whereas normal reticular-appearing mitochondria rarely colocalized with YFP-SMURF1 $\Delta$ C2. YFP-SMURF1(C699A) displayed the same subcellular staining pattern as wild-type YFP-SMURF1 (data not shown). Thus, SMURF1 colocalizes with damaged mitochondria in a C2 domain-independent manner.

We next determined whether the C2 domain of SMURF1 was required for the colocalization of damaged mitochondria with autophagosomes (Fig. 4e and Supplementary Figs 10 and 11). In cells expressing wild-type mCherry-SMURF1, mitochondria were mostly compacted around the nucleus, and numerous autophagosomes were observed surrounding structures that labelled positive for both mCherry-SMURF1 and the mitochondrial marker, Tom20 (also known as TOMM20). In contrast, in cells expressing mCherry-SMURF1 $\Delta$ C2, mCherry-SMURF1 $\Delta$ C2- and Tom20-positive mitochondria were rarely found inside autophagosomes. In many regions, GFP-LC3-positive linear or cup-shaped structures were observed near mCherry-SMURF1 $\Delta$ C2-positive mitochondria, but complete autophagosomes surrounding these mitochondria could not be detected. Thus, the C2 domain of SMURF1 is

not required for its targeting to damaged mitochondria, but is required for damaged mitochondria to be normally engulfed by autophagosomes. It is not yet known whether this requirement reflects a direct role for the C2 domain in binding to autophagosomal membrane phospholipids or is a more indirect consequence of other, as-yet-undescribed, effects of the C2 domain in mitophagy.

To investigate whether SMURF1 may function in selective autophagy *in vivo*, we performed electron microscopy analyses of cerebellum, liver and hearts of 10-month-old wild-type and *Smurf1*<sup>-/-</sup> mice<sup>21</sup>. In all three organs, *Smurf1*<sup>-/-</sup> mice showed an accumulation of abnormal mitochondria that were swollen, fragmented, and/or contained abnormal cristae (Fig. 4f). This phenotype is consistent with a defect in mitophagy and mitochondrial quality control; however, we cannot rule out unknown triggers of mitochondrial damage in these animals. In the livers of *Smurf1*<sup>-/-</sup> mice, mitochondria were spatially disorganized and surrounded by networks of dilated endoplasmic reticulum, perhaps reflecting a defect in mitochondrial targeting by isolation membranes (which are believed to originate from the endoplasmic reticulum<sup>1</sup>) and/or a defect in selective autophagy of the endoplasmic reticulum. There was a marked accumulation of lipid droplets in the livers of *Smurf1*<sup>-/-</sup> mice (Supplementary Fig. 12a), which may be consistent with selective degradation of lipid droplets by autophagy (lipophagy) in hepatocytes<sup>24</sup>. Furthermore, the granule cell layer of the cerebellum and cardiomyocytes of *Smurf1*<sup>-/-</sup> mice had increased numbers of p62 aggregates (Supplementary Fig. 12b). Unlike findings in brains and hearts of mice lacking core autophagy genes<sup>25</sup>, p62 aggregate accumulation in these tissues was not associated with ubiquitin accumulation. This is consistent with a role for SMURF1 in selective autophagy but not in the form of basal autophagy that is involved in protein quality control<sup>25</sup>.

Together, our data in *Smurf1*<sup>-/-</sup> MEFs and in *Smurf1*<sup>-/-</sup> mice suggest a crucial function for SMURF1 in selective autophagy, including in the autophagic targeting of genetically distinct viruses, in the autophagic targeting of mitochondria and, more speculatively, in the potential autophagic targeting of other cellular targets such as hepatic lipid droplets and endoplasmic reticulum. The mechanism by which SMURF1 functions in selective autophagy is independent of its E3 ubiquitin ligase activity, but rather involves its C2 membrane-targeting domain. We propose that the C2 domain of SMURF1 may participate in the delivery of selective autophagic substrates to the nascent autophagosome. Thus, SMURF1 has parallel functions in two distinct cellular degradation pathways, targeting specific proteins for degradation by the ubiquitin-proteasomal pathway<sup>15</sup> (via its E3 ubiquitin ligase activity) and targeting selective cargo for degradation by the autophagy pathway (via its C2 domain).

Our findings in *Smurf1*<sup>-/-</sup> MEFs and mice illustrate that our high-content image-based genome-wide screen successfully reveals novel candidate determinants of selective autophagy. More broadly, the identification of a set of 96 genes that may dually function in viral autophagy and mitophagy (but not in basal autophagy) suggests the existence of a common molecular network for targeting diverse unwanted cytoplasmic cargo to the lysosome. This network identification provides a basis for a more global understanding of the mechanisms involved in selective autophagy.

## METHODS SUMMARY

**High-content image-based genome-wide siRNA screen.** A genome-wide siRNA library (Dharmacon) containing 21,125 SMART pools was used for reverse transfection of HeLa/GFP-LC3 cells, followed by infection with SIN/mCherry.capsid virus, high content imaging using a Pathway855 automated microscope (BD Biosciences), quantitative image analysis, statistical analysis, and bioinformatic analysis as described in Supplementary Information. Primary hits were evaluated in three confirmation/secondary screens using the four individual siRNAs from each pool, including a screen for viral capsid/autophagosome colocalization, cell survival during SIN infection, and Parkin-induced mitophagy.

**Functional analyses of SMURF1.** See Supplementary Information.

Received 6 September 2010; accepted 7 September 2011.

Published online 23 October 2011.

- Levine, B., Mizushima, N. & Virgin, H. W. Autophagy in immunity and inflammation. *Nature* **469**, 323–335 (2011).
- Noda, N. N., Ohsumi, Y. & Inagaki, F. Atg8-family interacting motif crucial for selective autophagy. *FEBS Lett* **584**, 1379–1385 (2010).
- Wild, P. et al. Phosphorylation of the autophagy receptor optineurin restricts *Salmonella* growth. *Science* **333**, 228–233 (2011).
- Komatsu, M. & Ichimura, Y. Selective autophagy regulates various cellular functions. *Genes Cells* **15**, 923–933 (2010).
- Lipinski, M. M. et al. A genome-wide siRNA screen reveals multiple mTORC1 independent signaling pathways regulating autophagy under normal nutritional conditions. *Dev. Cell* **18**, 1041–1052 (2010).
- Orvedahl, A. O. et al. Autophagy protects against Sindbis virus infection of the central nervous system. *Cell Host Microbe* **7**, 115–127 (2010).
- Monastyrsky, I., Rieter, E., Klionsky, D. J. & Reggiori, F. Multiple roles of the cytoskeleton in autophagy. *Biol. Rev. Camb. Philos. Soc.* **84**, 431–448 (2009).
- Lee, J. Y. et al. HDAC6 controls autophagosome maturation essential for ubiquitin-selective quality-control autophagy. *EMBO J.* **29**, 969–980 (2010).
- Longatti, A. & Tooze, S. A. Vesicular trafficking and autophagosome formation. *Cell Death Differ.* **16**, 956–965 (2009).
- Nair, U. et al. SNARE proteins are required for macroautophagy. *Cell* **146**, 290–302 (2011).
- Behrends, C., Sowa, M. E., Gygi, S. P. & Harper, J. W. Network organization of the human autophagy system. *Nature* **466**, 68–76 (2010).
- Mizushima, N. The role of the Atg1/ULK1 complex in autophagy regulation. *Curr. Opin. Cell Biol.* **22**, 132–139 (2010).
- Narendra, D., Tanaka, A., Suen, D. F. & Youle, R. J. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *J. Cell Biol.* **183**, 795–803 (2008).
- Pagliarini, D. J. et al. A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123 (2008).
- Xing, L., Zhang, M. & Chen, D. Smurf control in bone cells. *J. Cell. Biochem.* **110**, 554–563 (2010).
- Taloczy, Z. et al. Regulation of starvation- and virus-induced autophagy by the eIF2α kinase signaling pathway. *Proc. Natl Acad. Sci. USA* **99**, 190–195 (2002).
- Taloczy, Z., Virgin, H. W. I. V. & Levine, B. PKR-dependent xenophagic degradation of herpes simplex virus type 1. *Autophagy* **2**, 24–29 (2006).
- Orvedahl, A. et al. HSV-1 ICP34.5 confers neurovirulence by targeting the Beclin 1 autophagy protein. *Cell Host Microbe* **1**, 23–35 (2007).
- Geisler, S. et al. PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. *Nature Cell Biol.* **12**, 119–131 (2010).
- Narendra, D., Kane, L. A., Hauser, D. N., Fearnley, I. M. & Youle, R. J. p62/SQSTM1 is required for Parkin-induced mitochondrial clustering but not mitophagy; VDAC1 is dispensable for both. *Autophagy* **6**, 1090–1106 (2010).
- Yamashita, M. et al. Ubiquitin ligase Smurf1 controls osteoblast activity and bone homeostasis by targeting MEK2 for degradation. *Cell* **121**, 101–113 (2005).
- Cho, W. & Stahelin, R. V. Membrane binding and subcellular targeting of C2 domains. *Biochim. Biophys. Acta* **1761**, 838–849 (2006).
- Lu, K. et al. Pivotal role of the C2 domain of the Smurf1 ubiquitin ligase in substrate selection. *J. Biol. Chem.* **286**, 16861–16870 (2011).
- Singh, R. et al. Autophagy regulates lipid metabolism. *Nature* **458**, 1131–1135 (2009).
- Mizushima, N. & Levine, B. Autophagy in mammalian development and differentiation. *Nature Cell Biol.* **12**, 823–830 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Vishwanath, S. Wei and B. Posner for assistance with high-throughput siRNA screening; W. Sun for information technology support; K. Scudder for assistance with image analysis algorithms; A. Diehl for expert medical illustration; V. Stollar, M. McDonald, R. Kuhn and R. Youle for helpful discussions and providing reagents; A. Bugde for assistance in the UTSW Live Cell Imaging Facility; and L. Mueller and T. Januszewski for assistance with electron microscopy. This work was supported by NIH grants AI109617 (B.L.), CA84254 (B.L.), UL1 RR024982 (G.X., Y.X.), AI062773 (R.J.X.), DK83756 (R.J.X.), DK086502 (R.J.X.) and DK043351 (R.J.X. and A.N.); NSF grant DMS-0907562 (G.X.); and the Center for Cancer Research, National Cancer Institute Intramural Research Program (Y.E.Z.).

**Author Contributions** A.O., R.S., M.N., M.R., J.L.W., Y.E.Z., K.L.-P., C.G. and B.L. designed the experiments. A.O., R.S., Z.Z. Q.S. and Y.T. performed the experiments. G.X., A.N., C.V.F., R.J.X. and Y.X. performed statistical and bioinformatic analyses. A.O., R.S. and B.L. wrote the manuscript. G.X. and A.N. contributed equally to the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to B.L. ([beth.levine@utsouthwestern.edu](mailto:beth.levine@utsouthwestern.edu)).

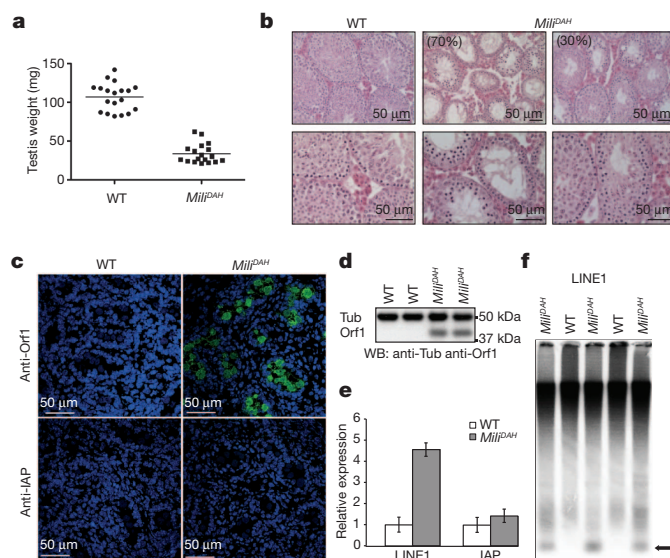
# The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements

Serena De Fazio<sup>1</sup>, Nenad Bartonicek<sup>2\*</sup>, Monica Di Giacomo<sup>1\*</sup>, Cei Abreu-Goodger<sup>2</sup>, Aditya Sankar<sup>1</sup>, Charlotta Funaya<sup>3</sup>, Claude Antony<sup>3</sup>, Pedro N. Moreira<sup>1</sup>, Anton J. Enright<sup>2</sup> & Dónal O'Carroll<sup>1</sup>

Piwi proteins and Piwi-interacting RNAs (piRNAs) have conserved functions in transposon silencing<sup>1</sup>. The murine Piwi proteins Mili and Miwi2 (also called Piwil2 and Piwil4, respectively) direct epigenetic LINE1 and intracisternal A particle transposon silencing during genome reprogramming in the embryonic male germ line<sup>2–4</sup>. Piwi proteins are proposed to be piRNA-guided endonucleases that initiate secondary piRNA biogenesis<sup>5–7</sup>; however, the actual contribution of their endonuclease activities to piRNA biogenesis and transposon silencing remain unknown. To investigate the role of Piwi-catalysed endonucleolytic activity, we engineered point mutations in mice that substitute the second aspartic acid to an alanine in the DDH catalytic triad of Mili and Miwi2, generating the *Mili*<sup>DAH</sup> and *Miwi2*<sup>DAH</sup> alleles, respectively. Analysis of Mili-bound piRNAs from homozygous *Mili*<sup>DAH</sup> fetal gonadocytes revealed a failure of transposon piRNA amplification, resulting in the marked reduction of piRNA bound within Miwi2 ribonuclear particles. We find that Mili-mediated piRNA amplification is selectively required for LINE1, but not intracisternal A particle, silencing. The defective piRNA pathway in *Mili*<sup>DAH</sup> mice results in spermatogenic failure and sterility. Surprisingly, homozygous *Miwi2*<sup>DAH</sup> mice are fertile, transposon silencing is established normally and no defects in secondary piRNA biogenesis are observed. In addition, the hallmarks of piRNA amplification are observed in Miwi2-deficient gonadocytes. We conclude that cycles of intra-Mili secondary piRNA biogenesis fuel piRNA amplification that is absolutely required for LINE1 silencing.

Transposable elements are mobile genetic elements that constitute a large fraction of eukaryotic genomes. The process of transposon silencing is of fundamental importance for genome integrity and germ cell development. Members of the Piwi subclade of the Argonaute proteins have conserved roles in transposon silencing and bind a class of small non-coding RNAs known as piRNAs that act as guides for targeting of the respective ribonuclear particles (RNPs)<sup>1</sup>. The Argonaute family is primarily defined by the presence of the Piwi domain that adopts a classical RNase H fold<sup>8</sup> with some Argonaute proteins being active small-RNA-guided endonucleases (slicers)<sup>9</sup>. Mechanisms of piRNA biogenesis are largely unclear, but two models summarize our current knowledge. First, primary processing of long single-stranded precursors by unknown nuclease(s) results in the generation of primary piRNAs. Thereafter, secondary biogenesis takes place via a 'ping-pong' cycle; that is, a feed-forward amplification loop wherein the slicer activities of two Piwi proteins take turns to generate secondary piRNAs using primary piRNAs as initial guides<sup>5,6</sup>. Originally described in *Drosophila*, where all three *Drosophila* Piwi proteins have proven endonuclease activity *in vitro*<sup>6,10</sup>, this secondary pathway model offers an explanation for transposon silencing as piRNA biogenesis consumes transposon transcripts<sup>5,6</sup>. Nevertheless, the role of Piwi-mediated slicing in secondary biogenesis or transposon silencing is not directly tested in any system.

In mice, transposable elements are epigenetically silenced throughout most of life with transposon repression being initially established during germ cell development<sup>11</sup>. The process of reprogramming in the fetal germ line initiates genome-wide CpG demethylation that erases both genomic imprints as well as retrotransposon promoter methylation<sup>11</sup>. In male mice the Piwi proteins Mili and Miwi2 are essential for *de novo* DNA methylation and repression of (L1) and intracisternal A particle (IAP) retrotransposons<sup>2–4</sup>. The mouse piRNA model proposes that transposon-derived primary Mili-bound sense piRNAs initiate a round of secondary processing by targeting complementary antisense transposon transcripts for cleavage by Mili's endonuclease activity<sup>7</sup>. The 3' cleavage fragment is then 3' processed, which results in the accumulation of secondary antisense piRNAs in Miwi2 that guide transposon DNA methylation. Furthermore, these



**Figure 1 | The endonuclease activity of Mili is required for spermatogenesis and L1 silencing.** **a**, Testicular atrophy in *Mili*<sup>DAH</sup> mice. Testicular weights of 3-month-old wild-type (WT) and *Mili*<sup>DAH</sup> mice are shown. **b**, Haematoxylin- and eosin-stained wild-type and *Mili*<sup>DAH</sup> testis sections from 3-month-old mice. The percentage of *Mili*<sup>DAH</sup> mice with the indicated phenotype is shown. **c**, Immunofluorescence using anti-L1 Orf1 and anti-IAP Gag antibodies (green) and DAPI-stained DNA (blue) on wild-type and *Mili*<sup>DAH</sup> E16.5 fetal testis sections are shown. **d**, Western blot (WB) using anti-L1 Orf1 and anti-tubulin antibodies on extracts from P10 wild-type and *Mili*<sup>DAH</sup> testes is shown. **e**, The expression levels of L1 and IAP were quantified by qRT-PCR from RNA derived from P10 wild-type and *Mili*<sup>DAH</sup> testes. Error bars indicate standard deviation from biological triplicates ( $n = 3$ ). **f**, Methylation-sensitive Southern blot on HpaII-digested DNA extracted from P10 wild-type and *Mili*<sup>DAH</sup> testis using a L1 promoter probe. The arrowhead indicates the identity of the methylation-sensitive fragment.

<sup>1</sup>European Molecular Biology Laboratory, Mouse Biology Unit, Via Ramarini 32, Monterotondo Scalo 00015, Italy. <sup>2</sup>European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, Cambridge CB10 1SD, UK. <sup>3</sup>European Molecular Biology Laboratory, EMBL Meyerhof Str. 1, 69117 Heidelberg, Germany.

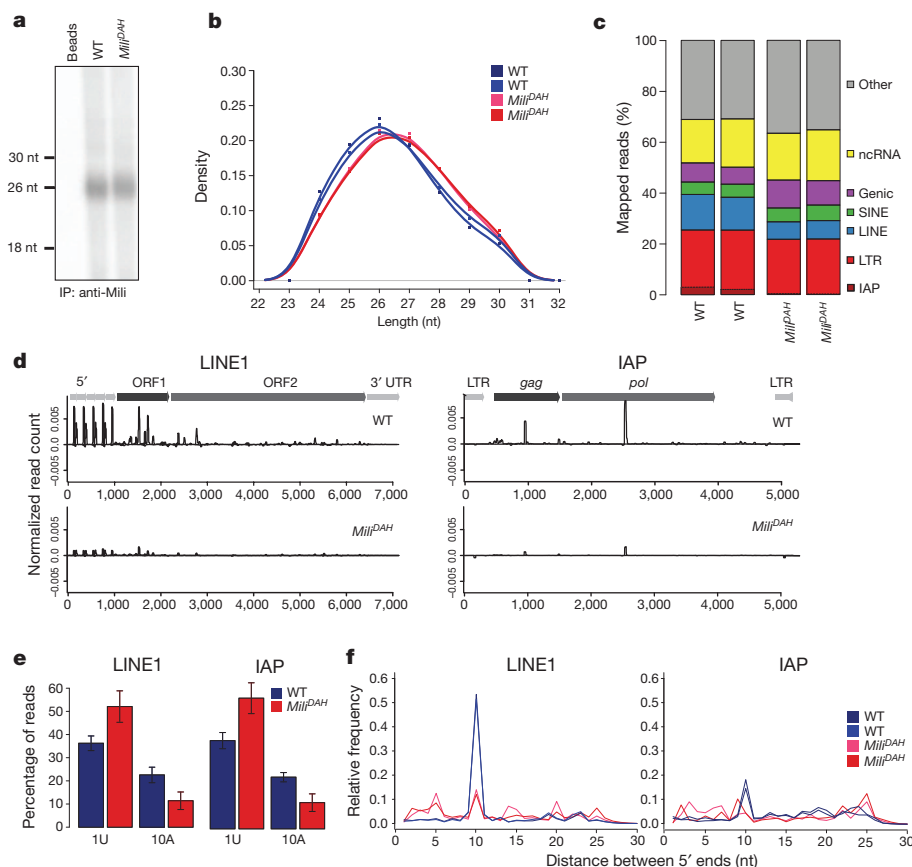
\*These authors contributed equally to this work.

antisense piRNA-programmed Miwi2 RNPs are also proposed to target transposon transcripts for cleavage followed by another round of secondary piRNA biogenesis that serves to amplify the primary transposon piRNAs<sup>7</sup>. Here we explore the contribution of the putative endonuclease activities of Mili and Miwi2 to the embryonic piRNA pathway and transposon silencing.

We decided to use the same genetic strategy that has recently been used to explore the *in vivo* endonuclease function of Argonaute 2 (Ago2)<sup>12–14</sup>. Mutation of the second aspartic acid to an alanine in the catalytic triad DDH (the Ago2<sup>D669A</sup> or Ago2<sup>DAH</sup> mutation/allele) abrogates Ago2's catalytic function without affecting protein stability and other functions of the protein *in vivo*<sup>13–16</sup>. We first generated knock-in *Mili*<sup>D813A/D813A</sup> or *Mili*<sup>DAH</sup> mice (Supplementary Fig. 1a–c). The point mutation did not have an impact on Mili protein expression (Supplementary Fig. 1d). Mili localizes to the inter-mitochondrial cement, a part of the germ-cell-specific perinuclear nuage<sup>17</sup>, and has been shown to interact with the Tudor domain protein Tdrd1 as well as the RNA helicases Mov10L1 and Mvh (refs 18–22). Co-localization of Mili with Tdrd1, Mov10L1 and Mvh was observed in fetal gonadocytes of wild-type and *Mili*<sup>DAH</sup> mice (Supplementary Fig. 2). Furthermore, interaction of Mili with Tdrd1 and Mov10L1 was confirmed in both wild-type and *Mili*<sup>DAH</sup> mice (Supplementary Table 1). In contrast to Mili, Tdrd1 or Mvh deficiency<sup>23,24</sup>, normal inter-mitochondrial cement was observed in *Mili*<sup>DAH</sup> gonadocytes (Supplementary Fig. 3). Therefore, the *Mili*<sup>DAH</sup> point mutation does not have an impact on

inter-mitochondrial cement integrity, expression levels or localization of Mili.

Deletion of Mili or Miwi2 leads to arrest in meiotic prophase<sup>3,25</sup>, probably attributable to secondary consequences of defective transposon DNA methylation and de-repression. All *Mili*<sup>DAH</sup> testes were atrophic (Fig. 1a), with most (70%) of *Mili*<sup>DAH</sup> mice presenting the same meiotic phenotype as *Mili*<sup>-/-</sup> mice. Early pachytene was the most advanced stage of germ cell development observed in these mice (Fig. 1b and Supplementary Fig. 4). The other 30% of *Mili*<sup>DAH</sup> mice presented a slightly milder phenotype. Tubules containing round and even elongating spermatids were observed, albeit with cells of aberrant morphology (Fig. 1b and Supplementary Fig. 5). All *Mili*<sup>DAH</sup> tubules were highly apoptotic (Supplementary Figs 4b and 5c). However, independent of the phenotype observed, all *Mili*<sup>DAH</sup> male mice are sterile (Supplementary Fig. 1e). The process of transposon *de novo* DNA methylation occurs during late gestation in mice, concluding in postnatal germ cells a few days after birth. De-repression of L1 elements was observed in *Mili*<sup>DAH</sup> embryonic day (E)16.5 gonadocytes and postnatal and adult germ cells (Fig. 1c–e and Supplementary Fig. 6a, b). With respect to the process of transposon silencing, no heterogeneity in adult *Mili*<sup>DAH</sup> mice was observed (Supplementary Fig. 6), raising the possibility that the distinct adult spermatogenic phenotypes may arise from a genetic background effect or that Mili and its associated endonuclease activity may function in processes beyond transposon silencing in meiotic or post-meiotic cells. Notably, in contrast to *Mili*<sup>-/-</sup>



**Figure 2 | piRNA amplification failure in *Mili*<sup>DAH</sup> mice.** **a**, Mili RNPs were immunoprecipitated from E16.5 fetal testis of the indicated genotypes. Associated piRNAs were visualized by 5' <sup>32</sup>P labelling, resolution on a 15% TBE urea gel and autoradiography. nt, nucleotide. **b**, Size profiles of cloned Mili-bound piRNAs from biological replicates of wild-type (blue) and *Mili*<sup>DAH</sup> (pink) E16.5 fetal gonadocytes. **c**, Genomic annotation of cloned Mili-bound piRNAs as indicated from pairs of biological replicates of wild-type and *Mili*<sup>DAH</sup> E16.5 fetal gonadocytes. LTR, long terminal repeat; ncRNA, non-coding RNA. **d**, Mapping of Mili-bound piRNAs to the consensus of L1 (left) and IAP (right) elements. Positive and negative values indicate sense and

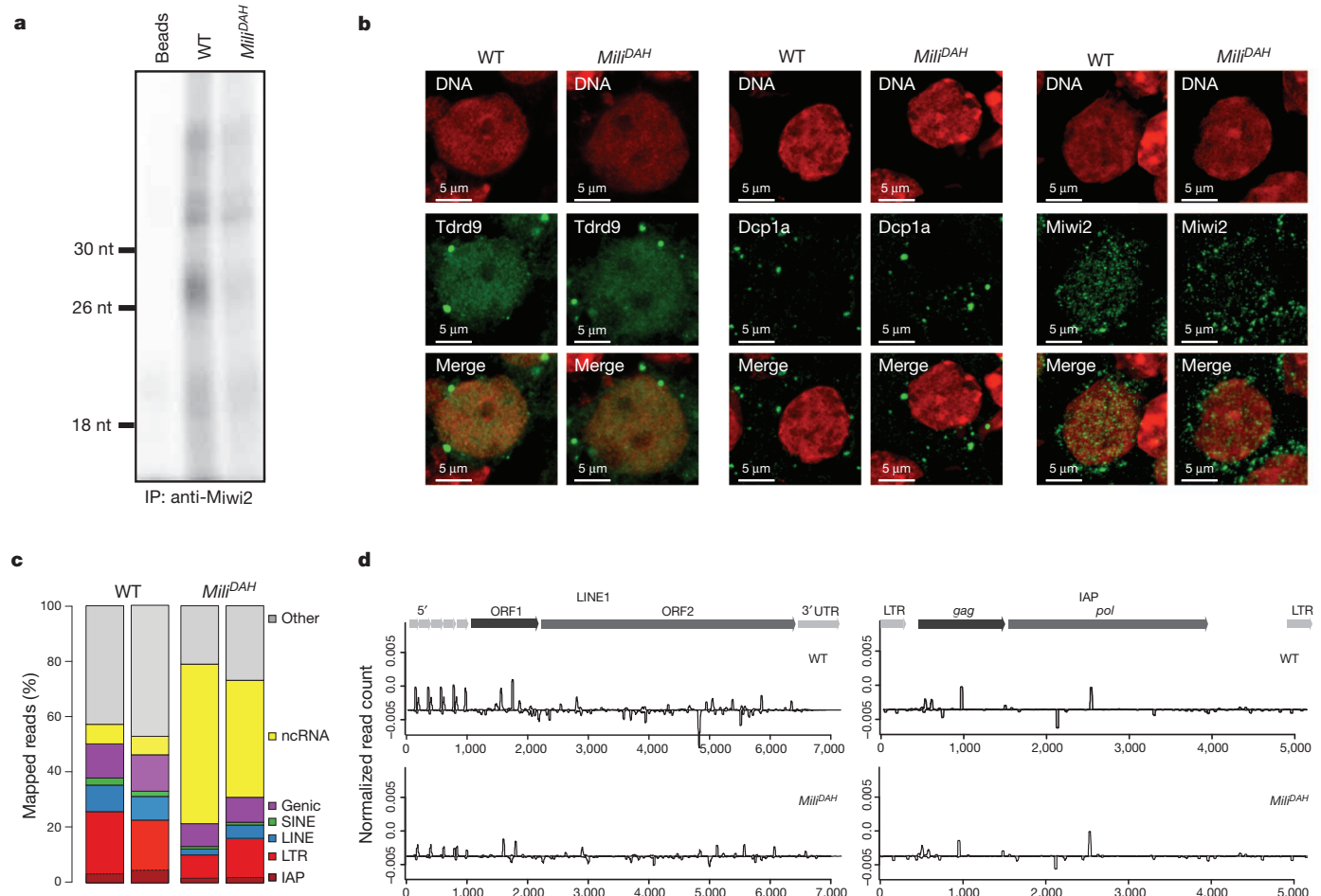
antisense piRNAs, respectively. Schematic representations of the respective elements are also shown (above). **e**, Percentage of piRNAs from wild-type and *Mili*<sup>DAH</sup> Mili RNPs with a U at the first position (1U) without A at position 10 and an A at position 10 (10A) without a U at position 1 are shown for L1 and IAP elements. Error bars indicate standard error of the mean from the biological duplicates ( $n = 2$ ). **f**, Ping-pong analysis of Mili-bound piRNAs from biological replicates of wild-type and *Mili*<sup>DAH</sup> E16.5 fetal gonadocytes. Relative frequency (y axis) of distances between 5' ends (x axis) between complementary piRNAs for both L1 and IAP elements is shown. nt, nucleotides.

mice<sup>2,4</sup>, IAPs are normally repressed in *Mili*<sup>DAH</sup> mice (Fig. 1c–e and Supplementary Fig. 6a, b). Accordingly, defective CpG DNA methylation of L1 but not IAP elements was observed in postnatal germ cells at time points after the completion of the *de novo* DNA methylation process (Fig. 1f and Supplementary Fig. 6c). Collectively, these data indicate that the endonuclease activity of *Mili* is absolutely required for L1 silencing.

We next analysed *Mili*-bound piRNAs to understand the impact of the *Mili*<sup>DAH</sup> mutation on the piRNA pathway. Equal quantities of piRNA were observed in *Mili* RNPs from *Mili*<sup>DAH</sup> fetal gonadocytes (Fig. 2a). piRNA libraries from biological duplicates of *Mili* RNPs from wild-type and *Mili*<sup>DAH</sup> fetal gonadocytes were subjected to deep sequencing. 20.3% of piRNAs are longer in *Mili*<sup>DAH</sup> versus wild-type gonadocytes (Fig. 2b). Among repetitive element piRNAs a 4.2- and 5.1-fold reduction in the fraction of L1 and IAP piRNAs, respectively, are observed in *Mili*<sup>DAH</sup> libraries (Fig. 2c). When we mapped the reads to the genome, no differences were observed in the genomic origin between wild-type and *Mili*<sup>DAH</sup> piRNAs (Supplementary Fig. 7), indicating that the endonuclease activity of *Mili* is not required for the selective expression of specific piRNA clusters. We next performed a qualitative analysis of L1 and IAP piRNAs. For this analysis, we only consider unique reads and not their individual depths within libraries. No differences in L1 and IAP piRNA formation were observed in the wild-type and *Mili*<sup>DAH</sup> libraries (Supplementary Fig. 8a). We therefore conclude that the *Mili*<sup>DAH</sup> mutation does not affect primary biogenesis of transposon or cluster-related piRNAs.

The precise endonucleolytic cleavage site of an Argonaute protein is at the nucleotide opposed to the tenth nucleotide from the 5' end of the guide small-RNA<sup>26,27</sup>. A 5'U (1U) bias is a characteristic of primary piRNAs, therefore when a primary piRNA directs cleavage of a target followed by 3' processing, the secondary piRNA will contain a bias for A at position 10 (refs 5, 6). In addition, there is an overlap of 10 nucleotides; this precise complementarity is detected when comparing distances between 5' ends of complementary piRNAs and is known as the ping-pong signature<sup>5,6</sup>. A second round of piRNA biogenesis fuelled by the secondary piRNA results in the amplification of the initiating primary piRNA. Analysis of *Mili*-bound piRNAs, now taking depth into consideration, revealed reduced levels of both L1 and IAP piRNAs in *Mili*<sup>DAH</sup> fetal gonadocytes (Fig. 2c, d). The increase in 1U and a decrease in 10A-containing piRNAs in *Mili*<sup>DAH</sup> libraries are indicative of the failure of piRNA amplification (Fig. 2e). Notably, the ping-pong signature is lost for both L1 and IAP piRNAs in *Mili*<sup>DAH</sup> mice (Fig. 2f). Collectively, these data are consistent with a function of the endonuclease activity of *Mili* in the initiation of secondary processing required for transposon piRNA amplification.

The lack of secondary processing in *Mili*<sup>DAH</sup> gonadocytes resulted in the stark reduction of piRNA in *Miwi2* RNPs (Fig. 3a). The additional cycles required to generate libraries from *Mili*<sup>DAH</sup> gonadocytes supported the observation of a significant decrease in piRNAs within *Miwi2* from *Mili*<sup>DAH</sup> gonadocytes (Supplementary Table 2). *Miwi2* localizes to cytoplasmic P-bodies known as piP bodies and to the nucleus<sup>7,17</sup>. Staining with the piP-body components Tdrd9 and

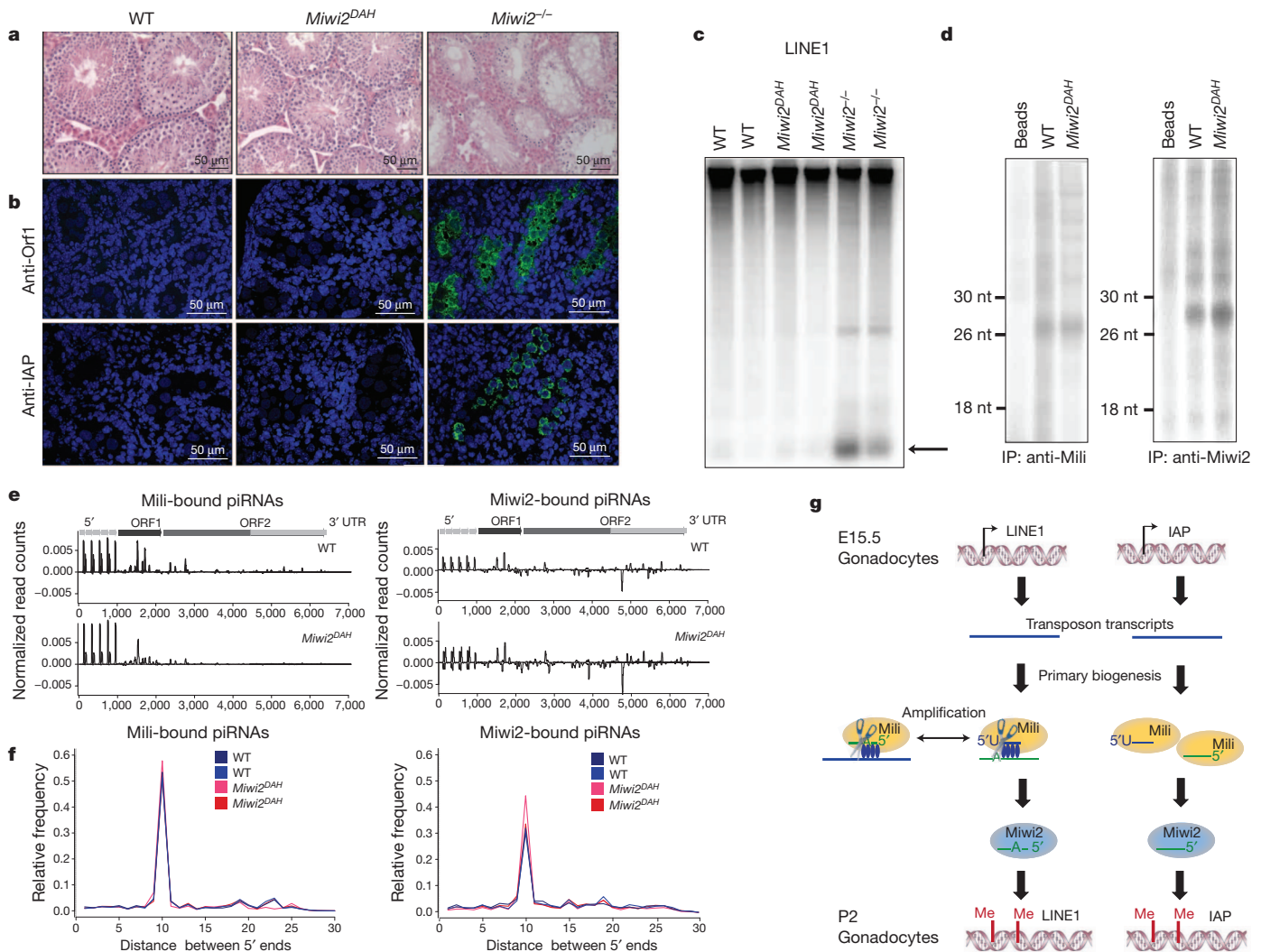


**Figure 3 | Marked reduction of *Miwi2*-bound piRNAs in *Mili*<sup>DAH</sup> mice.**

**a**, *Miwi2* RNPs were immunoprecipitated from E16.5 fetal testis of the indicated genotypes and piRNA represented as in Fig. 2a. **b**, Confocal projection images of indirect immunofluorescence with Tdrd9, Dcp1a and *Miwi2* antibodies (green) and DAPI-stained DNA (red) from E16.5 fetal testis of the indicated genotypes. **c**, Annotation of cloned *Miwi2*-bound piRNAs as

indicated from biological replicates of wild-type and *Mili*<sup>DAH</sup> E16.5 fetal gonadocytes. **d**, Mapping of the *Miwi2*-bound piRNAs to the consensus of L1 (left) and IAP (right) elements. Positive and negative values indicate sense and antisense piRNAs, respectively. Schematic representations of the respective elements are also shown.

Dcp1a revealed intact piP-body formation in *Mili*<sup>DAH</sup> gonadocytes (Fig. 3b). Furthermore, Miwi2 shows cytoplasmic foci and, consistent with the partial piRNA loading, retained nuclear localization in *Mili*<sup>DAH</sup> gonadocytes albeit with a lower staining intensity (Fig. 3b). Sequencing revealed a piRNA population of a normal size profile featuring an approximately 2-fold decrease in the percentage of both LINE- and IAP-associated piRNAs (Fig. 3c and Supplementary Fig. 9a). However, when mapped to L1 and IAP consensus sequences, the Miwi2-bound piRNAs from *Mili*<sup>DAH</sup> mice showed a pattern of piRNA formation equivalent to wild type (Fig. 3d and Supplementary Fig. 8b). Whereas Miwi2 bound sense and antisense transposon piRNAs in almost equal measure in both wild-type and *Mili*<sup>DAH</sup> gonadocytes (Supplementary Table 2), as previously described<sup>4,7</sup>, the amplitude of the antisense piRNA peaks derived across L1 was diminished compared to those previously observed<sup>4,7</sup>. This difference probably reflects experimental variation due to the antibody used or recent advances in library generation/sequencing. No difference in the genomic origin of cluster-related piRNAs was observed (Supplementary Fig. 9b).



**Figure 4 | Normal spermatogenesis and transposon silencing in *Miwi2*<sup>DAH</sup> mice.** **a**, Haematoxylin- and eosin-stained wild-type, *Miwi2*<sup>DAH</sup> and *Miwi2*<sup>-/-</sup> testis sections from 3-month-old mice. **b**, Immunofluorescence using anti-L1 Orf1 and anti-IAP Gag antibodies (green) and DAPI-stained DNA (blue) on wild-type, *Miwi2*<sup>DAH</sup> and *Miwi2*<sup>-/-</sup> E16.5 fetal testis sections are shown. **c**, Methylation-sensitive Southern blot on HpaII-digested DNA extracted from P10 wild-type, *Miwi2*<sup>DAH</sup> and *Miwi2*<sup>-/-</sup> testis using a L1 promoter probe is shown. The arrow indicates the identity of the methylation-sensitive fragment. **d**, Mili (left) and Miwi2 (right) RNPs were immunoprecipitated from E16.5 fetal testis of the indicated genotypes shown as in Fig. 2a. **e**, Mapping of the

Therefore, the endonuclease activity of Mili does not alter the identity of piRNAs that reside within Miwi2 RNPs, merely their quantity.

The fact that Miwi2 is seeded with the appropriate but severely reduced quantities of piRNAs in *Mili*<sup>DAH</sup> mice indicates that the potential endonuclease activity of Miwi2 cannot compensate for the loss of Mili's slicer activity. To understand the contribution of Miwi2's putative endonuclease activity to piRNA amplification and transposon silencing, we generated *Miwi2*<sup>D718A/D718A</sup> (*Miwi2*<sup>DAH</sup>) and *Miwi2*<sup>-/-</sup> mice (Supplementary Figs 10 and 11). Surprisingly, *Miwi2*<sup>DAH</sup> mice are fertile with no defects observed in testis morphology (Fig. 4a). Furthermore, in contrast to *Miwi2*<sup>-/-</sup> fetal gonadocytes<sup>3,4</sup>, both L1 and IAP are normally repressed (Fig. 4b, c). In addition to transposon silencing, Miwi2 is specifically required for the maintenance of spermatogonial stem cells<sup>3</sup>. No loss of germ cells is observed in *Miwi2*<sup>DAH</sup> aged mice (9–12 months) (Supplementary Fig. 12). Therefore, the *Miwi2*<sup>DAH</sup> mutation does not have an impact on any known physiological function of Miwi2. Accordingly, both Mili and Miwi2 are normally loaded with piRNA in *Miwi2*<sup>DAH</sup> fetal gonadocytes

Mili-bound (left) and Miwi2-bound (right) piRNAs to the consensus of L1. Positive and negative values indicate sense and antisense piRNAs, respectively. Schematic representation of L1 is shown (above). **f**, Ping-pong analysis of Mili- and Miwi2-bound piRNAs from biological replicates of wild-type and *Miwi2*<sup>DAH</sup> E16.5 fetal gonadocytes are shown. The frequency of the distance between 5' ends of complementary piRNAs for L1 is presented as in Fig. 2e. **g**, Model. L1 element silencing is dependent upon Mili's endonuclease activity for piRNA amplification. IAP silencing is dependent upon Mili and Miwi2 but independent of piRNA amplification. The small blue and green lines represent sense and antisense piRNAs, respectively.

(Fig. 4d). No qualitative or quantitative differences in piRNA origin or amplification were observed in Mili- or Miwi2-bound piRNAs from *Miwi2<sup>DAH</sup>* fetal gonadocytes (Fig. 4e, f and Supplementary Figs 13–15). The lack of physiological and molecular phenotype in *Miwi2<sup>DAH</sup>* mice argues against the possibility that Miwi2-mediated cleavage events functionally contribute to either slicing of nuclear transposon transcripts or piRNA amplification necessary for the establishment of transposon silencing. In support of the latter conclusion, we find piRNA amplification as judged by the existence of a robust ping-pong signature in Mili-bound piRNAs from *Miwi2<sup>-/-</sup>* fetal gonadocytes (Supplementary Fig. 16).

We propose a revised model of mammalian embryonic piRNA function, whereby Mili's endonuclease activity initiates secondary piRNA processing with an intra-Mili ping-pong cycle fuelling piRNA amplification (Fig. 4g). Mili's endonuclease activity is specifically required for the expansion of L1 and IAP piRNAs within Mili RNPs and the normal accumulation of all classes of piRNAs within Miwi2 RNPs. The defective piRNA pathway in *Mili<sup>DAH</sup>* mice results in the failure to repress specifically L1, revealing distinct silencing requirements for the respective transposons. The establishment of L1- but not IAP-silencing is strictly dependent upon Mili's endonuclease activity and secondary piRNA biogenesis (Fig. 4g). The sufficiency of primary piRNA processing in *Mili<sup>DAH</sup>* mice to direct Miwi2-mediated IAP repression illustrates fundamental differences in piRNA dosage required for silencing of the respective transposons. This difference may reflect their genomic burden—L1 occupies approximately 19% of the mouse genome whereas IAPs account for only 0.2% (ref. 28). Therefore, amplification of the L1 piRNA pool may be essential to program sufficient Miwi2 to target the prolific L1 element. In conclusion, we show that a single Piwi endonuclease supports piRNA amplification and distinct transposon silencing. It remains to be seen if these are conserved features of piRNA-mediated transposon silencing.

## METHODS SUMMARY

**Histology.** Testes were fixed in Bouin's fixative overnight at 4 °C temperature, paraffin embedded and sectioned at 6–8-µm thickness. Sections were then stained (haematoxylin and eosin or periodic acid Schiff) by routine methods.

**Immunofluorescence.** For immunofluorescence, E16.5 fetal testes were freshly embedded in OCT, 6 µm sections cut and fixed in 4% paraformaldehyde. For Mili co-localization with Mvh, Tdrd1 and Mov10L1, sections were boiled twice for 5 min in 10 mM pH 6 sodium citrate solution for antigen retrieval. Sections were blocked for 30 min at room temperature in 10% normal donkey serum, 2% BSA and 0.1 M glycine. Primary antibody incubation was done overnight at 4 °C in the blocking buffer. Anti-rabbit Alexa-488-conjugated (1:1,000; Invitrogen) and anti-mouse Alexa-546-conjugated (1:1,000; Invitrogen) antibodies were used as secondary antibodies. DAPI (3 µg µl<sup>-1</sup>) (Sigma) was used to stain DNA. For Miwi2, L1 Orf1 and IAP staining, paraformaldehyde-fixed sections were permeabilized with 0.1% Triton X-100 for 10 min and then blocked in TBS-T 10% normal donkey serum and processed as above. A Leica TCS SP5 confocal microscope was used to acquire images.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 31 January; accepted 8 September 2011.

Published online 23 October 2011.

- Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nature Rev. Genet.* **10**, 94–108 (2009).
- Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G. J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**, 744–747 (2007).
- Carmell, M. A. *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **12**, 503–514 (2007).
- Kuramochi-Miyagawa, S. *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–917 (2008).
- Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103 (2007).
- Gunawardane, L. S. *et al.* A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* **315**, 1587–1590 (2007).

- Aravin, A. A. *et al.* A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* **31**, 785–799 (2008).
- Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434–1437 (2004).
- Patel, D. J. *et al.* Structural biology of RNA silencing and its functional implications. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 81–93 (2006).
- Saito, K. *et al.* Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev.* **20**, 2214–2222 (2006).
- Bestor, T. H. & Bourc'his, D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb. Symp. Quant. Biol.* **69**, 381–388 (2004).
- Cheloufi, S., Dos Santos, C. O., Chong, M. M. & Hannon, G. J. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**, 584–589 (2010).
- Cifuentes, D. *et al.* A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**, 1694–1698 (2010).
- O'Carroll, D. *et al.* A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes Dev.* **21**, 1999–2004 (2007).
- Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–1441 (2004).
- Maniatakis, E. & Mourelatos, Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev.* **19**, 2979–2990 (2005).
- Aravin, A. A. *et al.* Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. *PLoS Genet.* **5**, e1000764 (2009).
- Kojima, K. *et al.* Associations between PIWI proteins and TDRD1/MTR-1 are critical for integrated subcellular localization in murine male germ cells. *Genes Cells* **14**, 1155–1165 (2009).
- Reuter, M. *et al.* Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nature Struct. Mol. Biol.* **16**, 639–646 (2009).
- Vagin, V. V. *et al.* Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.* **23**, 1749–1762 (2009).
- Wang, J., Saxe, J. P., Tanaka, T., Chuma, S. & Lin, H. Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. *Curr. Biol.* **19**, 640–644 (2009).
- Zheng, K. *et al.* Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. *Proc. Natl Acad. Sci. USA* **107**, 11841–11846 (2010).
- Chuma, S. *et al.* Tdrd1/Mtr-1, a tudor-related gene, is essential for male germ-cell differentiation and nuage/germinal granule formation in mice. *Proc. Natl Acad. Sci. USA* **103**, 15894–15899 (2006).
- Kuramochi-Miyagawa, S. *et al.* MVH in piRNA processing and gene silencing of retrotransposons. *Genes Dev.* **24**, 887–892 (2010).
- Kuramochi-Miyagawa, S. *et al.* Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* **131**, 839–849 (2004).
- Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200 (2001).
- Martinez, J. & Tuschl, T. RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes Dev.* **18**, 975–980 (2004).
- Mandal, P. K. & Kazanian, H. H. Jr. SnapShot: Vertebrate transposons. *Cell* **135**, 192–192 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are grateful to R. Pillai, B. Cullen, S. Martin, S. Chuma and J. Lykke-Andersen for antibodies used in this study. This study was technically supported by of EMBL's genomic core facility. We are very grateful to V. Benes, R. Pillai and S. van Dongen for advice. We are grateful to M. Reuter for assistance with the preparation of immunoprecipitations for mass spectrometry. This study was technically supported by EMBL Monterotondo's FACS and Microscopy core facilities. We are grateful to A. Wutz for A9 ES cells. We also acknowledge the services of J. Rientjes from Monash University's Gene Recombining Facility. We are also very grateful to C. Kutter and D. Odom for advice on small RNA library generation.

**Author Contributions** S.D.F. contributed to the design, execution and analysis of the majority of experiments on *Mili<sup>DAH</sup>* and *Miwi2<sup>DAH</sup>* mice. N.B. performed the bioinformatic analysis presented in the manuscript with initial assistance from C.A.-G. M.D.G. analysed the spermatogenic defects as well as undertook the co-localization studies in the respective mouse strains. A.S. performed the bisulphite sequencing experiments. C.F. and C.A. performed the electron microscopy experiments. P.N.M. established the 8-cell embryo ES cell injection procedure. A.J.E. supervised the bioinformatic analysis. D.O.C. conceived and supervised this study and wrote the final version of the manuscript.

**Author Information** All raw sequencing data are deposited in ArrayExpress (accession number E-MTAB-73) and European Nucleotide Archive (ERP000778). The *Mili<sup>DAH</sup>*, *Miwi2<sup>DAH</sup>* and *Miwi2* null (*Miwi2<sup>-/-</sup>*) alleles have been deposited at EMMA (<http://www.emmanet.org/>) and will be freely available on a non-collaborative basis. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.O.C. ([ocarroll@embl.it](mailto:ocarroll@embl.it)).

## METHODS

**Mouse strains.** *Mili*<sup>DAH</sup> allele: the Piwi domain of Mili is encoded in exons 20 to 23. Mili aspartic acid 813, the second aspartic acid of the catalytic triad, is encoded in exon 21. To generate the *Mili*<sup>D813A</sup> allele we replaced wild-type exon 21 with a mutant exon where the aspartic acid 813 codon is mutated to encode an alanine. A targeting construct was recombineered that contains homology arms and a *frt* flanked neomycin (*neo*) cassette 3' of exon 21 that contains the *Mili*<sup>D813A</sup> mutation. Southern blotting of the individual ES-cell-derived clone genomic *SacI*-digested DNA with a 3' external probe was used to identify homologous recombinants. A 9.2-kb DNA fragment corresponds to the wild-type *Mili* locus; integration of the *neo*<sup>-</sup>*frt* flanked cassette 3' of exon 21 introduces an additional *SacI* site, thus decreasing the size of the *SacI* DNA fragment recognized to 8.0 kb in the targeted allele. FLP-mediated recombination and excision of the *neo*<sup>-</sup>*frt* flanked cassette results in a 6-kb *SacI* DNA fragment recognized by the external 3' probe, which is diagnostic of the *Mili*<sup>DAH</sup> allele.

*Miwi2*<sup>DAH</sup> allele: the Piwi domain of Miwi2 is encoded in exons 14–19. Miwi2 aspartic acid 761, the second aspartic acid of the catalytic triad, is encoded in exon 17. To generate the *Miwi2*<sup>D761A</sup> allele we replaced wild-type exon 17 with a mutant exon where the aspartic acid 761 codon is mutated to encode an alanine. A targeting construct was recombineered that contains homology arms and a *frt* flanked neomycin (*neo*) cassette 3' of exon 17 that contains the *Miwi2*<sup>D761A</sup> mutation. Southern blotting of the individual ES-cell-derived clone genomic *BsrGI*-digested DNA with an external 5' probe was used to identify homologous recombinants. An 11.8-kb DNA fragment corresponds to the wild-type *Miwi2* locus; integration of the *neo*<sup>-</sup>*frt* flanked cassette 3' of exon 17 introduces an additional *BsrGI* site, thus decreasing the size of the *BsrGI* DNA fragment recognized to 8.4 kb. FLP-mediated recombination and excision of the *neo*<sup>-</sup>*frt* flanked cassette results in a 6.4-kb *BsrGI* DNA fragment recognized by the external 3' probe, which is diagnostic of the *Miwi2*<sup>DAH</sup> allele.

*Miwi2* null allele: for the *Miwi2*<sup>-</sup> allele, we flanked exon 17 with *loxP* sites. Cre-mediated deletion of exon 17 results in out-of-frame splicing between exon 16 and 18, resulting in stop codons before the last exon, targeting the mutant transcript for nonsense-mediated decay<sup>29</sup>. Should a fraction of the mutant transcript escape nonsense-mediated decay, a truncated protein would be made that lacks most of the Piwi domain and thus would probably be non-functional. To generate this allele, a targeting construct was generated that contains the same homology arms as the *Mili*<sup>D813A</sup> construct, an *frt* flanked *neo* cassette with a *loxP* site 3' of exon 17 and a second 5' *loxP* site. Southern blotting of the individual ES-cell-derived clone genomic *KpnI*-digested DNA with an external 5' probe was used to identify homologous recombinants. A 10.1-kb DNA fragment corresponds to the wild-type *Miwi2* locus; integration of the *neo*<sup>-</sup>*frt2-loxP* cassette 3' of exon 17 introduces an additional *KpnI* site, thus decreasing the size of the *KpnI* DNA fragment recognized to 8.0 kb. Cre-mediated recombination and excision of exon 17 and *neo*<sup>-</sup> *loxP* flanked cassette results in a 5.3-kb *KpnI* DNA fragment recognized by the external 3' probe, which is diagnostic of the *Miwi2* null (*Miwi2*<sup>-</sup>) allele.

The *Miwi2* targeting constructs and the *Mili* targeting construct were electroporated into IB10<sup>30</sup> and A9 ES cells, respectively. A9 ES cells are derived from hybrid embryos resulting from a 129/Sv male by C57BL/6 female cross (A. Wutz, manuscript in preparation). Southern blotting as described above of the individual ES-cell-derived clones was used to identify homologous recombinants. IB10-targeted ES cells were used to generate chimaeras for the respective *Miwi2* targeted alleles by standard blastocyst injections. A9-targeted ES cells were injected into C57BL/6 8-cell-stage embryos for the generation of fully ES-cell-derived mice following a procedure similar to one described previously<sup>31</sup> but using a Piezo Impact Unit (PMM150FU, Prime Tech) rather than a laser to puncture the zona pellucida of the host embryo. The *Mili*<sup>DAH</sup> and *Miwi2*<sup>DAH</sup>-targeted mice were then crossed to the FLP-expressing transgenic mice (FLPeR)<sup>32</sup> to remove the *frt* flanked *neo*<sup>-</sup> cassette, resulting in the generation of *Mili*<sup>DAH</sup> and *Miwi2*<sup>DAH</sup> alleles, respectively. Mice heterozygous for the *loxP* flanked *Miwi2*-targeted allele were crossed to Deleter Cre<sup>33</sup> to generate the *Miwi2* null (*Miwi2*<sup>-</sup>) allele. The mice analysed in this study were on a mixed C57BL/6 and 129 genetic background.

All of the mice were bred and maintained in EMBL Mouse Biology Unit, Monterotondo in accordance with current Italian legislation (Art. 9, 27 January 1992, number 116) under license from the Italian health ministry.

**Antibodies.** Rabbit polyclonal antibodies against mouse *Miwi2* were generated using the same epitope as described previously<sup>4</sup> and used for immunoprecipitation and immunofluorescence (1:200) experiments. A mouse monoclonal antibody against Mili was obtained from R. Pillai and used for immunoprecipitation and immunofluorescence (1:1,000) experiments. The following antibodies were used at the indicated dilutions for immunofluorescence: anti-Orf1 L1 (S. Martin; 1:250), anti-IAP Gag (B. Cullen; 1:500), anti-Mvh (Abcam (ab13840); 1:200),

anti-Mov10L1 (R. Pillai; 1:200), anti-Tdrd1 (R. Pillai; 1:200), anti-Tdrd9 (S. Chuma; 1:200) and anti-Dcpl1a (J. Lykke-Andersen; 1:500).

**Histology.** Testes were fixed in Bouin's fixative overnight at 4 °C temperature, paraffin embedded and sectioned at 6–8-μm thickness. Sections were then stained (haematoxylin and eosin or periodic acid Schiff) by routine methods.

**Immunofluorescence.** For immunofluorescence, E16.5 fetal testes were freshly embedded in OCT, 6 μm sections cut and fixed in 4% paraformaldehyde. For Mili co-localization with Mvh, Tdrd1 and Mov10L1, sections were boiled twice for 5 min in 10 mM pH 6 sodium citrate solution for antigen retrieval. Sections were blocked for 30 min at room temperature in 10% normal donkey serum, 2% BSA and 0.1 M glycine. Primary antibody incubation was done overnight at 4 °C in the blocking buffer. Anti-rabbit Alexa-488-conjugated (1:1,000; Invitrogen) and anti-mouse Alexa-546-conjugated (1:1,000; Invitrogen) antibodies were used as secondary antibodies. DAPI (3 μg μl<sup>-1</sup>) (Sigma) was used to stain DNA. For Miwi2, L1 Orf1 and IAP staining, paraformaldehyde-fixed sections were permeabilized with 0.1% Triton X-100 for 10 min and then blocked in TBS-T 10% normal donkey serum and processed as above.

A Leica TCS SP5 confocal microscope was used to acquire images.

**Detection of apoptotic cells.** Detection of apoptotic cells was performed on paraformaldehyde-fixed paraffin-embedded testis section using the *in situ* cell death detection kit (Roche) and developed with DAB substrate (Roche). Sections were also stained with haematoxylin.

**Electron microscopy.** Fetal testes were fixed in 2.5% glutaraldehyde in 50 mM cacodylate buffer supplemented with 2% sucrose, 50 mM KCl, 2.6 mM CaCl<sub>2</sub> and 2.6 mM MgCl<sub>2</sub> for 30 min at 4 °C and rinsed in 50 mM cacodylate buffer. Samples were incubated in 2% osmium in 50 mM cacodylate buffer for 40 min on ice, rinsed in water and incubated in 0.5% uranylacetate in water for 30 min on ice. The contrast enhancement procedure was followed by a stepwise dehydration in ethanol, up to 100% ethanol and infiltration in EPON (Roth) for embedding. Polymerization was done at 60 °C. Ultrathin sections (60–70-nm thickness) of the testis were obtained using an ultramicrotome (Leica Microsystems) and the sections were mounted on formvar-coated slot grids, and contrasted with uranylacetate and lead citrate. The sections were then viewed in a CM120 biotwin electron microscope (FEI) operating at 100 kV. Digital acquisitions were made with a Keen View CCD camera (Soft Imaging System).

**Germ cell isolation.** The previously described *Oct4-GFP* (ref. 34) allele was crossed into *Mili*<sup>DAH</sup> and *Miwi2*<sup>DAH</sup> mice to label germ cells with eGFP. To isolate E16.5 and postnatal day 7 germ cells, single cell suspensions of testis were obtained by two-step enzymatic digestion and GFP-positive FACS sorted.

**CpG methylation analysis.** Methylation-sensitive Southern blotting and bisulphite methylation analysis were performed as described<sup>4</sup>.

**Mass spectrometry.** Mili was immunoprecipitated as described<sup>19</sup>, resolved on gel and the entire lane with the exception of the immunoglobulin fragments was subjected to liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) on a LTQ Orbitrap Velos (Thermo Fisher Scientific) instrument.

**RT-qPCR analysis.** Total RNA was isolated from 10 day post-partum testis using Trizol according to the manufacturer's instruction and treated with Turbo DNase RNase-free (Ambion). cDNA synthesis was performed with SuperScript III Reverse Transcriptase (Invitrogen) with Random Hexamers (Invitrogen). Quantitative PCR was carried out by using SYBR Green I Master mix (Roche) on LightCycler 480 system (Roche). Three animals for each genotype were examined and assays were always done in triplicate. Primers for qPCR were used as described<sup>3</sup>.

**Small RNA library generation.** Mili and *Miwi2* RNPs were immunoprecipitated as described<sup>19</sup>. Small RNA libraries were generated as described previously<sup>35</sup> but using adaptors suitable for sequencing on the Illumina platform.

**Data analysis.** Sequencing data were processed from the FASTQ format and analysed using R/Bioconductor<sup>36</sup>. Barcode sequences were resolved by sample with no mismatches allowed, 5'/3' adaptors were stripped, sequences were filtered for low-complexity regions and finally size-selected for reads between 24–30 nt. Processed reads for each sample were mapped against the mouse genome (NCBI m37) using Bowtie 0.12.5 (ref. 37) allowing for two mismatches and requesting all matching sites. For reads mapping to multiple distinct loci, only the first 100 were reported according to match score. Mapped reads were categorized according to genomic annotations from Ensembl<sup>38</sup> Mouse v58 (LTR, LINE, SINE, genic, non-coding RNA). Reads not mapping to any recorded genomic element were classed as 'other'. Counts for reads mapping to multiple loci were divided by the total number of loci. In the specific cases of LINE and IAP elements, all reads were also mapped against representative canonical sequences obtained from GenBank<sup>39</sup> (M13002.1, EU183301.1) using Bowtie allowing for up to three mismatches. The number of repeat mapping reads was divided by repeat length and the number of genome mapping reads per billion processed reads. Genome-wide visualizations of read mapping to the mouse genome were obtained using Circos<sup>40</sup> plots and were

scaled between each set of four samples according to library size (see figure legend). Line-track *y*-axis maxima are set to one-tenth of the highest peak from the library with most reads. Heat-map minima were set to one-half the average read count across all bins and maxima to one-half the maximum read count across all bins. Heat-map scaling across colour-space was performed using the 'scale\_log\_base = 5' parameter of Circos to sample colour space better.

For ping-pong analysis, only reads mapping to repeat elements were considered. For each pair of sense/antisense overlapping reads, the distance between their 5' ends was recorded and counts were represented as relative frequencies within samples for each repeat element.

29. Wen, J. & Brogna, S. Nonsense-mediated mRNA decay. *Biochem. Soc. Trans.* **36**, 514–516 (2008).
30. Robanus-Maandag, E. *et al.* p107 is a suppressor of retinoblastoma development in pRb-deficient mice. *Genes Dev.* **12**, 1599–1609 (1998).
31. Poueymirou, W. T. *et al.* F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. *Nature Biotechnol.* **25**, 91–99 (2007).
32. Farley, F. W., Soriano, P., Steffen, L. S. & Dymecki, S. M. Widespread recombinase expression using FLPeR (flipper) mice. *Genesis* **28**, 106–110 (2000).
33. Schwenk, F., Baron, U. & Rajewsky, K. A *cre*-transgenic mouse strain for the ubiquitous deletion of *loxP*-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23**, 5080–5081 (1995).
34. Yoshimizu, T. *et al.* Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41**, 675–684 (1999).
35. Hafner, M. *et al.* Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **44**, 3–12 (2008).
36. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
37. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
38. Flicek, P. *et al.* Ensembl 2011. *Nucleic Acids Res.* **39**, D800–D806 (2011).
39. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. GenBank. *Nucleic Acids Res.* **36**, D25–D30 (2008).
40. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

# Ascaris suum draft genome

Aaron R. Jex<sup>1\*</sup>, Shiping Liu<sup>2\*</sup>, Bo Li<sup>2\*</sup>, Neil D. Young<sup>1\*</sup>, Ross S. Hall<sup>1</sup>, Yingrui Li<sup>2</sup>, Linfeng Yang<sup>2</sup>, Na Zeng<sup>2</sup>, Xun Xu<sup>2</sup>, Zijun Xiong<sup>2</sup>, Fangyuan Chen<sup>2</sup>, Xuan Wu<sup>2</sup>, Guojie Zhang<sup>2</sup>, Xiaodong Fang<sup>2</sup>, Yi Kang<sup>2</sup>, Garry A. Anderson<sup>1</sup>, Todd W. Harris<sup>3</sup>, Bronwyn E. Campbell<sup>1</sup>, Johnny Vlaminck<sup>4</sup>, Tao Wang<sup>4</sup>, Cinzia Cantacessi<sup>1</sup>, Erich M. Schwarz<sup>5</sup>, Shoba Ranganathan<sup>6</sup>, Peter Geldhof<sup>4</sup>, Peter Nejsun<sup>7</sup>, Paul W. Sternberg<sup>5</sup>, Huanming Yang<sup>2</sup>, Jun Wang<sup>2</sup>, Jian Wang<sup>2</sup> & Robin B. Gasser<sup>1</sup>

Parasitic diseases have a devastating, long-term impact on human health, welfare and food production worldwide. More than two billion people are infected with geohelminths, including the roundworms *Ascaris* (common roundworm), *Necator* and *Ancylostoma* (hookworms), and *Trichuris* (whipworm), mainly in developing or impoverished nations of Asia, Africa and Latin America<sup>1</sup>. In humans, the diseases caused by these parasites result in about 135,000 deaths annually, with a global burden comparable with that of malaria or tuberculosis in disability-adjusted life years<sup>1</sup>. *Ascaris* alone infects around 1.2 billion people and, in children, causes nutritional deficiency, impaired physical and cognitive development and, in severe cases, death<sup>2</sup>. *Ascaris* also causes major production losses in pigs owing to reduced growth, failure to thrive and mortality<sup>2</sup>. The *Ascaris*-swine model makes it possible to study the parasite, its relationship with the host, and ascariasis at the molecular level. To enable such molecular studies, we report the 273 megabase draft genome of *Ascaris suum* and compare it with other nematode genomes. This genome has low repeat content (4.4%) and encodes about 18,500 protein-coding genes. Notably, the *A. suum* secretome (about 750 molecules) is rich in peptidases linked to the penetration and degradation of host tissues, and an assemblage of molecules likely to modulate or evade host immune responses. This genome provides a comprehensive resource to the scientific community and underpins the development of new and urgently needed interventions (drugs, vaccines and diagnostic tests) against ascariasis and other nematodiasis.

We sequenced the *A. suum* genome at ~80-fold coverage (Supplementary Fig. 1), producing a final draft assembly of 272,782,664 base pairs (bp) (N50 = 407 kilobases, kb; N90 = 80 kb; 1,618 contigs of

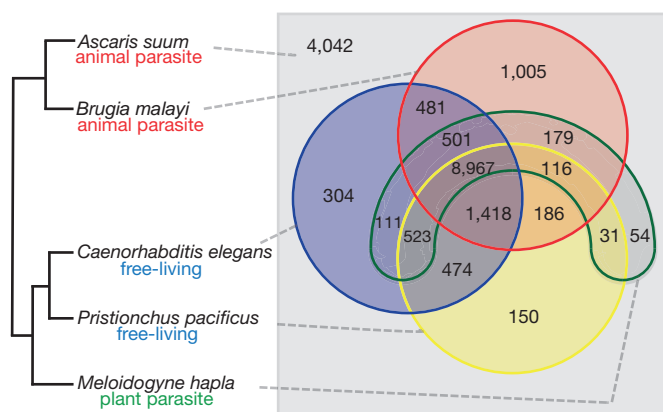
>2 kb) (Table 1) with a mean GC-content of 37.9%. This genome has few repetitive sequences (about 4.4% of the total assembly) relative to that reported for other metazoan genomes sequenced to date<sup>3–6</sup>, probably as a result of chromatin diminution<sup>7</sup>. We identified 424 distinct retrotransposon sequences (see Supplementary Tables 1–3) representing at least 22 families (8 long terminal repeats (LTRs), 12 long interspersed elements (LINEs) and 2 short interspersed elements (SINEs)), with *Gypsy*, *Pao* and *Copia* classes predominating for LTRs ( $n = 97$ , 85 and 60, respectively) and CR1, L1, and reverse transcriptase encoding RTE-RTE classes predominating for non-LTRs ( $n = 29$ , 28 and 21, respectively). We also identified eight families of DNA transposons (91 distinct sequences in total), of which *MuDr*, *En-Spm* and *Merlin* ( $n = 12$ , 9 and 8, respectively) predominated. We predicted 18,542 genes (14,783 supported by transcriptomic data), with a mean total length of 6.5 kb, exon length of 153 bp and a mean of 6.4 exons per gene (see Supplementary Fig. 2). Compared with the nematodes (roundworms) *Caenorhabditis elegans*<sup>3</sup>, *Pristionchus pacificus*<sup>8</sup>, *Brugia malayi*<sup>9</sup> or *Meloidogyne hapla*<sup>10</sup>, overall, the *A. suum* genes are significantly longer (see Supplementary Table 2), relating primarily to expansions of intronic regions (mean 1.1 kb).

Most (78.2%) of the predicted *A. suum* genes (Fig. 1) have a homologue (BLASTp cut-off  $\leq 10^{-5}$ ) either in *C. elegans* ( $n = 12,779$ ; 68.9%), *B. malayi* (12,853; 69.3%), *M. hapla* (10,482; 56.5%) or *P. pacificus* (11,865; 64.0%), with 8,967 being homologous among all species examined, and 4,042 (21.8%) being 'unique' to *A. suum* (see Fig. 1). Of the genes with homology to *C. elegans* or *B. malayi*, ~50%

**Table 1 | Features of the *Ascaris suum* draft genome**

Estimated genome size in megabases	309
Total number of base pairs within assembled scaffolds	272,782,664
N50 length in bp; total number >2 kb in length	407,899; 1,618
N90 length in bp; total number >N90 length	80,017; 748
GC content of whole genome (%)	37.9
Repetitive sequences (%)	4.4
Proportion of genome that is coding (exonic; including introns) (%)	5.9; 44.2
Number of putative coding genes	18,542
Gene size (mean bp)	6,536
Average coding domain length (mean bp)	983
Average exon number per gene (mean)	6
Gene exon length (mean bp)	153
Gene intron length (mean bp)	1,081
GC content in coding regions (%)	45
Number of transfer RNAs	255

N50 means 50% of all nucleotides in the assembly are within contigs of  $\geq 408$  kb. N90 means 90% of all nucleotides in the assembly are within contigs of  $\geq 80$  kb. Genome size estimated on the basis of  $k$ -mer (see online-only Methods) frequency.



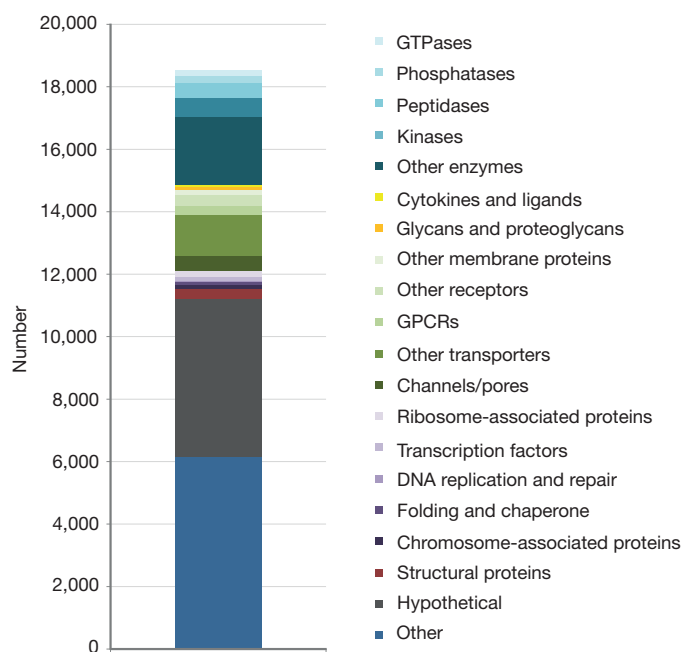
**Figure 1 | Venn diagram summarizing the overlapping homology between the *Ascaris suum* gene set and those of other nematodes.** Grey box (right) represents genes unique to *A. suum*, relative to *Brugia malayi* (red circle), *Caenorhabditis elegans* (blue circle), *Meloidogyne hapla* (green arc) and/or *Pristionchus pacificus* (yellow circle). The phylogram (left) displays the evolutionary relationships currently proposed among the nematodes.

<sup>1</sup>Faculty of Veterinary Science, The University of Melbourne, Parkville, Victoria 3010, Australia. <sup>2</sup>BGI-Shenzhen, Shenzhen, 518083, China. <sup>3</sup>Ontario Institute for Cancer Research, MaRS Centre, South Tower, 101 College Street, Suite 800, Toronto, Ontario, M5G 0A3, Canada. <sup>4</sup>Laboratory of Parasitology and Parasitic Diseases, University of Ghent, Merelbeke B-9820, Belgium. <sup>5</sup>California Institute of Technology, Pasadena, California, 91125, USA. <sup>6</sup>Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia. <sup>7</sup>Faculty of Life Sciences, University of Copenhagen, Copenhagen DK-2200, Denmark.

\*These authors contributed equally to this work.

and 44%, respectively, were determined to represent one-to-one orthologues<sup>11</sup> (see Supplementary Data 1). For these orthologues (on scaffolds exceeding one megabase, 1 Mb, in size), we explored synteny for *A. suum* and *B. malayi* by pairwise comparison with *C. elegans* (see Supplementary Data 1). The findings show that interchromosomal gene rearrangements in *A. suum* are relatively rare and occurred less frequently in *A. suum* than in *B. malayi*<sup>9</sup> relative to *C. elegans* since their evolutionary divergence<sup>12</sup>. In contrast, intrachromosomal rearrangements were relatively common and comparable in frequency to those inferred for *B. malayi*<sup>9</sup>. Overall synteny was significantly higher between *A. suum* and *B. malayi* (~15%) than between either species and *C. elegans* (~3%), which is consistent with current knowledge of the evolutionary relationships among these three species<sup>12</sup>. Interestingly, of these *C. elegans* orthologous genes, 532 and 483 were exclusive to the current assemblies of the *A. suum* and *B. malayi* genomes, respectively (Supplementary Data 2). Although there were no homology matches between these two exclusive subsets of orthologues, they shared striking similarity in functional ontology (biological process), being linked predominantly to growth, reproduction, development and/or morphogenesis. There is clear evidence of plasticity in the germline of metazoans<sup>13</sup>, with cases of products from non-homologous genes in different species having analogous function(s). Therefore, we hypothesize that these two unique gene subsets relate to differences in reproductive biology (oviparity versus viviparity) and life history (direct versus indirect) between *A. suum* and *B. malayi*. Clearly, this proposal warrants testing and functional validation in *C. elegans* and/or in *Ascaris*.

Of the entire *A. suum* gene set, 2,370 genes had an orthologue (BLASTp cut-off  $\leq 10^{-5}$ ) belonging to one of 279 known biological (KEGG; see online-only Methods) pathways (Supplementary Data 3). Mapping to pathways in *C. elegans* indicated a full complement of molecules; by inference, the vast majority (95%) of the *A. suum* euchromatin is represented in the present genomic assembly, an inference that is supported by our transcriptomic data (Supplementary Tables 4 and 5). We were able to assign possible functions (such as for enzymes, receptors, channels and transporters; Supplementary Fig. 3, Supplementary Table 6 and Supplementary Data 4) to 13,503 (72.8%) of the genes predicted for *A. suum* (Fig. 2). For these genes, we predicted 456 peptidases belonging to five major classes (aspartic, cysteine, metallo-, serine and threonine), with the metallo- ( $n = 184$ ;



**Figure 2** | The major protein classes representing the *Ascaris suum* gene set.

41.0%) and serine proteases ( $n = 132$ : 30.0%) predominating (Supplementary Data 4). Notably, the secreted peptidases (such as the M12 'astacins', the S9 and S33 serine proteases, and the C1 and C2 cysteine proteases) are abundantly represented, and have key roles in tissue invasion and degradation (for example, during migration and/or feeding) and/or immune evasion/modulation in many parasites<sup>14,15</sup>.

In addition, we identified 609 kinases and 257 phosphatases, respectively (Supplementary Data 4). All major classes of kinases are represented, with the tyrosine (TK:  $n = 94$ ), casein (CK1:  $n = 83$ ), CMGC ( $n = 67$ ) and CAMK ( $n = 54$ ) being most abundant in *A. suum*. The phosphatome includes 17 receptor and 68 conventional tyrosine, 64 serine/threonine and 39 dual-specificity phosphatases. On the basis of homology with molecules in *C. elegans*, 169 GTPases are encoded in the *A. suum* genome, including 135 small GTPases (Ras superfamily) representing the Rab ( $n = 36$ ), Ras ( $n = 35$ ; plus 8 Ras-like), Rho ( $n = 17$ ; plus 9 Rho-like) or Ran ( $n = 6$ ) subfamilies. Examples of these homologues include *eft-1*, *fzo-1*, *glo-1* and *rho-1*, which have essential roles in embryonic, larval and/or reproductive development (see [www.wormbase.org](http://www.wormbase.org)).

Given their key roles, many of these enzymes are proposed as targets for anti-parasitic compounds and/or vaccines<sup>16–18</sup>. Equally, the range of receptor and channel proteins identified here are interesting because many common anthelmintics bind such targets<sup>19</sup>. Here, we predicted 279 G protein-coupled receptors (GPCRs) for *A. suum* and 477 channel or pore proteins (Supplementary Data 4), including 272 voltage-gated and 98 ligand-gated ion channels. Many voltage-gated ion channels are known targets for nematocidal drugs, such as macrocyclic lactones (for example, ivermectin) and levamisole, and an aminoacetonitrile derivative, monepantel, is the most recent example of a highly effective nematocide that binds to a ligand-gated ion channel<sup>19</sup>. Importantly, in the *A. suum* gene set, we found a homologue (*acr-23*) of the *C. elegans* monepantel receptor<sup>19</sup>, suggesting that this drug may kill *A. suum*. In addition, we detected 462 transporters (for example, small molecule porter proteins), of which the major facilitator ( $n = 155$ ), cation symporter ( $n = 71$ ) and resistance-nodulation-cell division ( $n = 56$ ) superfamilies were most abundant (Supplementary Data 4).

Excretory/secretory (E/S) peptides are central to understanding parasite–host interactions. We predicted the secretome of *A. suum* to comprise 775 proteins with diverse functions (Supplementary Data 5). Notable among them are 68 secreted proteases, including 20 SC clan serine proteases (S9 and S33 families), 18 MA clan metallo-proteases (M10, M12 and M41 families) and 5 CA/CD clan cysteine proteases (C1 and C13 families); see <http://merops.sanger.ac.uk/> for clan definitions.

Secreted proteases have known roles in host-tissue degradation, required for feeding, tissue-penetration and/or larval migration for a range of helminths<sup>14</sup>, including *Ascaris*<sup>2</sup>. In addition, they are involved in inducing and modulating host immune responses against parasitic helminths<sup>15</sup>, which are often Th2-biased<sup>20</sup>. From the current understanding of these responses<sup>15</sup>, we compiled a comprehensive list of *A. suum* E/S proteins homologous to helminth-secreted peptides with important immunogenic or immunomodulatory roles in host animals (Supplementary Table 7 and Supplementary Data 6). Such homologues represent about half of the predicted *A. suum* secretome. Most abundant among them are O-linked glycosylated proteins ( $n = 300$ ), many of which are heavily targeted by immunoglobulin (Ig) M antibodies and bound by various pattern recognition receptors associated with host dendritic cells responsible for the induction of a Th2 immune response<sup>15</sup>.

Other members of the *A. suum* secretome are predicted to direct or evade immune responses. These peptides include a close homologue of the E/S-62 leucyl aminopeptidase of the filarioid nematode *Acanthocheilonema viteae*, which has been shown to inhibit B-cell, T-cell and mast cell proliferation/responses, promote an alternative activation of the host macrophages, through the inhibition of the Toll-like receptor signalling pathway, and induce a Th2 response through

the inhibition of IL-12p70 production by dendritic cells<sup>15</sup>. Additional, immunomodulatory molecules predicted for *A. suum* (BLASTp cut-off  $\leq 10^{-5}$ ) include homologues of another B-cell inhibitor (that is, the *B. malayi* cystatin CPI-2), several TGF- $\beta$  and macrophage initiation factor mimics, numerous neutrophil inhibitors, various oxidoreductases, and five close homologues of platelet anti-inflammatory factor  $\alpha$  (ref. 15). Some *A. suum* E/S peptides are predicted to be involved in immune evasion; for instance, some mask parasite antigens by mimicking host molecules (such as several C-type lectins with close homology to vertebrate macrophage mannose or CD23 (low affinity IgE receptors<sup>15</sup>).

Taken together, these data indicate that *A. suum* has a large arsenal of E/S proteins that are likely to be involved directly in manipulating, blocking and/or evading immune responses in the host. Understanding the immunomolecular interplay between *A. suum* and its host, early in infection, particularly during hepatopulmonary migration, should pave the way for designing prophylactic interventions, such as vaccination.

*Ascaris* larvae undertake an extensive migration through their host's body before they establish as adults in the small intestine. Following the ingestion of infective eggs and their gastric passage, third-stage larvae (L3s)<sup>21</sup> hatch from eggs in the gut and penetrate the intestinal wall; they then undergo, via the bloodstream, an arduous hepatopulmonary migration. The complexity of this migration coincides with important developmental changes in the nematode<sup>2</sup>. Clearly, this migration requires tightly regulated transcriptional changes in the parasite. We explored this aspect by characterizing the transcription profiles of infective L3s (from eggs), L3s from the liver or lungs of the host, and fourth-stage larvae (L4s) from the small intestine (Supplementary Fig. 4, Supplementary Data 7). Notable among genes enriched during larval migration are various secreted peptidases linked to tissue-penetration and degradation during feeding and/or migration<sup>14</sup>, including three C1/C2, five M1, eight M12, fourteen S9 and five S33 clan members. Considering the complex nature of larval migration, a key role for molecules associated with chemosensory pathways is highly likely. Such molecules have been studied extensively in *C. elegans*<sup>22</sup>, with numerous homologues being identified here in larval transcripts (Supplementary Data 7). With few exceptions, all of these homologues relate to olfactory chemosensation of volatile compounds (for example, alcohols, aldehydes or ketones), suggesting that the olfactory detection of molecular gradients is central to the navigation of *A. suum* larvae during migration. Lastly, considering the substantial host attack against migrating *Ascaris* larvae, E/S proteins probably play crucial roles in immune modulation and/or evasion during hepatopulmonary migration. Many such genes, including *Bm-alt-1*, *Bm-cpi-2* and *mif-4*, are highly transcribed in *A. suum* larvae (see Supplementary Data 7), particularly in migrating L3s.

Because of the large size of the adult nematode (10–15 cm), we were able to explore transcription in the musculature and reproductive tracts of adult male and female *A. suum* individuals as well (Supplementary Fig. 5 and Supplementary Data 8). Among the male-enriched transcripts is a range of genes associated specifically with sperm and/or spermatogenesis, including *fer-1*, *spe-4*, *spe-6*, *spe-9*,

*spe-10*, *spe-15* and *spe-41*, *alg-4* and *msp-57* (see [www.wormbase.org](http://www.wormbase.org)). Notable among the female-enriched transcripts is a large variety of genes associated with oogenesis/egg-laying (such as *cat-1*, *unc-54*, *cbd-1* and *pqn-74*), vulval development (such as *noah-1*, *nhr-25*, *cog-1* and *pax-3*) and/or embryogenesis (such as *cam-1* and *unc-6*; see [www.wormbase.org](http://www.wormbase.org)). Although the functions of these genes have been explored in *C. elegans* (primarily a hermaphroditic nematode), this detailed insight into the tissue-specific transcription for a dioecious nematode is a major advance.

Analyses of these RNA-seq data revealed 163,777 single nucleotide polymorphisms (SNPs) in coding regions of the *A. suum* genome; 61% of them were synonymous, 7% non-synonymous and <0.1% termination codons (Supplementary Data 9). Some of the most variable genes in *A. suum* encoded ribosomal proteins ( $n = 44$ ), translation initiation factor (*tif*) eIF-3 subunits 3 and 5, *tif* TFIIH subunit H2 and *tif* IF-2, galectin-4 and galectin-9, the latter two of which are probably linked to immune evasion<sup>15</sup> and may indicate that antigenic variation is among the many strategies apparent in *Ascaris* to combat the host immune response. Interestingly, the high nucleotide variability linked to the key elements of translation machinery did not relate to a bias in synonymous SNPs, suggesting that many mutations accumulate in particular 'hotspots' and/or are tolerated, but do not compromise either the structure or the function of this machinery. The least variable genes encoded various (druggable)<sup>16,17</sup> serine/threonine phosphatases ( $n = 17$ ) as well as numerous receptors, channels and transporters, for which there was an unusually strong bias towards synonymous SNPs, reinforcing their potential as intervention targets.

Given our present reliance on a small number of drugs (for example, piperazine, pyrantel, albendazole and mebendazole) for the treatment of ascariasis, their repeated or excessive use might lead to resistance in *Ascaris* populations to some or all of these compounds<sup>23</sup>. As few new anthelmintics (that is, aminoacetylnitriles<sup>19</sup> and cyclooctodepsipeptides<sup>24</sup>) have been discovered in the past two decades using traditional screening methods, an effective, alternative means of drug discovery is urgently needed<sup>23</sup>. Genome-guided drug target or drug discovery has major potential to complement conventional screening and re-purposing. The goal of genome-guided analysis is to identify genes or molecules whose inactivation by one or more drugs will selectively kill parasites but not harm their host.

Because most parasitic nematodes are difficult to produce or maintain outside of their host, or to subject to gene-specific silencing by RNAi<sup>23</sup> or morpholinos<sup>25,26</sup>, direct functional assessment of essentiality (that is, they are needed for the nematode's survival) is not yet practical. However, essentiality can be inferred from functional information for model organisms (for example, lethality in *C. elegans* and *D. melanogaster*)<sup>27</sup>, and this approach has indeed yielded effective targets for nematocides<sup>16</sup>. In *Ascaris*, we identified 629 proteins (Supplementary Data 10) with essential homologues in *C. elegans* and *D. melanogaster* (linked to lethal phenotypes upon gene perturbation). Among these are 87 channels or transporters (including 44 voltage-gated ion channels), which represent protein classes most successfully targeted for anthelmintic compounds, including macrocyclic lactones, levamisoles and aminoacetonitrile

**Table 2 | Druggable candidates in the *Ascaris suum* draft genome**

Protein or chokepoint	Subtype (number of molecules)	Total number
GTPase	Small GTPase (22); Ras (13); Rab (5); Rho (3); Ras-like (1); others (2)	46
Kinase	TK (8); AGC (3); CAMK (2); TKL (2); STE (1); other (1)	17
Peptidase	A22A (5); M14B (3); M12B (2); M67A (1); C14A (1); C50 (1); M12A (1); M13 (1); T01A (1); C46 (1); S33 (1)	19
Phosphatase	STP (28); cPTP (4); DSP (3)	35
Transporters and channels	Channels and pores (30); primary active transporters (24); incompletely characterized transport system proteins (22); accessory factors involved in transport (5); electrochemical potential-driven transporters (5); group translocators (1)	87
'Lethal' chokepoints	CDP-diacylglycerol-inositol 3-phosphatidyltransferase	1
	G protein-coupled receptor kinase 5	1
	Phosphoribosylformylglycinamide synthase	1
	Inosine-5'-monophosphate dehydrogenase	1
	Phospho-N-acetylmuramoyl pentapeptide transferase	1

Candidates were inferred from essentiality prediction and metabolic chokepoint analysis.

derivatives<sup>19,28</sup>. Also notable are 46 GTPases, 35 phosphatases (including PP1 and PP2A homologues, as targets for norcantharidin analogues)<sup>16</sup>, 17 kinases and 19 peptidases (Table 2).

In addition to essentiality-based prediction, an alternative strategy has been to infer enzymatic chokepoints intrinsic to the complete metabolome of a parasite<sup>29</sup>. Such chokepoints are defined as enzymatic reactions that uniquely produce and/or consume a molecular compound, using the strategy that the disruption of such enzymes would lead to the toxic build-up (that is, for unique substrates) or starvation (that is, for unique products) of metabolites within cells. Pathway analysis identified 225 likely chokepoints linked to genes predicted to be essential in *A. suum* (Supplementary Data 10). We gave the highest priority to targets predicted from single-copy genes in the *A. suum* genome, reasoning that lower allelic variability would exist within populations and would thus be less likely to give rise to drug resistance.

Using this strategy, we identified five high-priority drug targets for *A. suum* (see Table 2 and Supplementary Data 10) that, given their conservation with *C. elegans* and *D. melanogaster*, are likely to be relevant in relation to many other parasitic helminths. Conspicuous among them is IMP dehydrogenase (GMP reductase), which has a variety of inhibitors (for example, mycophenolic acid analogues<sup>30</sup>) that could be tested for ascaricidal effects. Clearly, the druggable genome of *Ascaris* now provides a solid basis for rational drug design, aimed at controlling parasitic nematodes of major socioeconomic impact worldwide.

In conclusion, we have characterized the genome of *A. suum*, a major parasite of one of the world's most important food animals (pig) and the closest relative of *A. lumbricoides*, which infects about 1.2 billion people globally<sup>1,2</sup>. Intriguingly, the present *A. suum* draft genome exhibits unusually low repeat content and lacks Tas2 transposons<sup>7</sup>. These characteristics probably relate to the chromatin diminution described previously for some ascaridoids<sup>7</sup>, indicating that our assembly represents the somatic genome of this parasite. The precise mechanism governing this diminution is not yet understood. Although the chromatin lost during this process is not fully characterized, there appears to be a significant loss in repeat content<sup>7</sup>, consistent with the present assembly. Notably, the present gene set inferred for *A. suum* includes *fert-1* and *rpS19G*, which, although originally proposed to be germline-specific<sup>7</sup>, were transcribed in all adult libraries sequenced here. This finding suggests that the genomic content lost during diminution might vary among individuals or tissues, and is a stimulus to investigate chromatin diminution between and among individual cells (that is, sperm or eggs), stages and tissue types of *A. suum*. Importantly, the present study, showing that a high-quality genomic assembly can be achieved using an approach based on whole-genome amplification, provides unique prospects for exploring diminution in detail, using the present genome as a reference.

In addition, our sequencing effort has characterized a broad range of key classes of molecules of major relevance to understanding the molecular biology of *A. suum* and the exquisite complexities of the host–parasite interplay on an immunobiological level. This work paves the way for future fundamental molecular explorations and the design of new methods for the treatment and control of one of the world's most important parasitic nematodes. This focus is now crucial, given the major impact of *Ascaris* and other soil-transmitted helminths, which affect billions of people and animals worldwide. Although these parasites are seriously neglected, genomic and post-genomic approaches provide new hope for the discovery of intervention strategies, with major implications for improving global health.

## METHODS SUMMARY

We sequenced the genome of *A. suum* using Illumina technology from genomic DNA from the reproductive tract of a single adult female. From six paired-end sequencing libraries (insert sizes: 0.17 kb to 10 kb; see Supplementary Tables 1 and 2), we generated 39 Gb of useable short-read sequence data, equating to ~80-fold coverage of the 273-Mb genome. We assembled the short reads, constructed

scaffolds in a step-by-step manner, and then closed intra-scaffold gaps<sup>5</sup>. Transposable elements, non-coding RNAs and the protein-coding gene set were inferred using a combination of predictive modelling and a homology-based approach. Orthology and synteny analyses were conducted using established methods<sup>9,11</sup>. We sequenced messenger RNA from infective L3s (from eggs), migrating L3s from the liver or lungs of the host, and L4s from the small intestine, as well as muscle and reproductive tissues from adult male and female *A. suum*, and used these data to aid gene predictions, define SNPs and explore key molecules associated with larval migration, reproduction and development. All proteins predicted from the gene set were annotated using databases for conserved protein domains, gene ontology annotations and model organisms (that is, *Caenorhabditis elegans*, *Drosophila melanogaster* and *Mus musculus*). Essentiality and drug target predictions were conducted using established or in-house methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 16 June; accepted 12 September 2011.**

**Published online 26 October 2011.**

- Hotez, P. J., Fenwick, A., Savioli, L. & Molyneux, D. H. Rescuing the bottom billion through control of neglected tropical diseases. *Lancet* **373**, 1570–1575 (2009).
- Crompton, D. W. *Ascaris* and ascariasis. *Adv. Parasitol.* **48**, 285–375 (2001).
- The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018 (1998).
- Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
- Li, R. *et al.* The sequence and *de novo* assembly of the giant panda genome. *Nature* **463**, 311–317 (2010).
- Mitrev, M. *et al.* The draft genome of the parasitic nematode *Trichinella spiralis*. *Nature Genet.* **43**, 228–235 (2011).
- Müller, F. & Tobler, H. Chromatin diminution in the parasitic nematodes *Ascaris suum* and *Parascaris univalens*. *Int. J. Parasitol.* **30**, 391–399 (2000).
- Dieterich, C. *et al.* The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nature Genet.* **40**, 1193–1198 (2008).
- Ghedini, E. *et al.* Draft genome of the filarial nematode parasite *Brugia malayi*. *Science* **317**, 1756–1760 (2007).
- Opperman, C. H. *et al.* Sequence and genetic map of *Meloidogyne hapla*: a compact nematode genome for plant parasitism. *Proc. Natl Acad. Sci. USA* **105**, 14802–14807 (2008).
- Kuzniar, A., van Ham, R. C. H. J., Pongor, S. & Leunissen, J. A. M. The quest for orthologs: finding the corresponding gene across genomes. *Trends Genet.* **24**, 539–551 (2008).
- Blaxter, M. L. *et al.* A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**, 71–75 (1998).
- Ewen-Campen, B., Schwager, E. E. & Extavour, C. G. The molecular machinery of germ line specification. *Mol. Reprod. Dev.* **77**, 3–18 (2010).
- McKerrow, J. H., Caffrey, C., Kelly, B., Loke, P. & Sajid, M. Proteases in parasitic diseases. *Annu. Rev. Pathol.* **1**, 497–536 (2006).
- Hewitson, J. P., Grainger, J. R. & Maizels, R. M. Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Mol. Biochem. Parasitol.* **167**, 1–11 (2009).
- Campbell, B. E. *et al.* Norcantharidin analogues with nematocidal activity in *Haemonchus contortus*. *Bioorg. Med. Chem. Lett.* **21**, 3277–3281 (2011).
- Campbell, B. E., Hofmann, A., McCluskey, A. & Gasser, R. B. Serine/threonine phosphatases in socioeconomically important parasitic nematodes—prospects as novel drug targets? *Biotechnol. Adv.* **29**, 28–39 (2011).
- Renslo, A. R. & McKerrow, J. H. Drug discovery and development for neglected parasitic diseases. *Nature Chem. Biol.* **2**, 701–710 (2006).
- Kaminsky, R. *et al.* A new class of anthelmintics effective against drug-resistant nematodes. *Nature* **452**, 176–180 (2008).
- Maizels, R. M. & Yazdanbakhsh, M. Immune regulation by helminth parasites: cellular and molecular mechanisms. *Nature Rev. Immunol.* **3**, 733–744 (2003).
- Geenen, P. L. *et al.* The morphogenesis of *Ascaris suum* to the infective third-stage larvae within the egg. *J. Parasitol.* **85**, 616–622 (1999).
- Bargmann, C. I. Chemosensation in *C. elegans* in *Wormbook* (ed. The *C. elegans* Research Community) (2006); <http://www.wormbook.org>.
- Keiser, J. & Utzinger, J. The drugs we have and the drugs we need against major helminth infections. *Adv. Parasitol.* **73**, 197–230 (2010).
- Harder, A. *et al.* Cyclooctadepsipeptides—an anthelmintically active class of compounds exhibiting a novel mode of action. *Int. J. Antimicrob. Agents* **22**, 318–331 (2003).
- Heasman, J. Morpholino oligos: making sense of antisense? *Dev. Biol.* **243**, 209–214 (2002).
- Geldhof, P. *et al.* RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects. *Parasitology* **134**, 609–619 (2007).
- Lee, I. *et al.* A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. *Nature Genet.* **40**, 181–188 (2008).
- Campbell, W. C., Fisher, M. H., Stapley, E. O., Albers-Schonberg, G. & Jacob, T. A. Ivermectin: a potent new antiparasitic agent. *Science* **221**, 823–828 (1983).
- Berriman, M. *et al.* The genome of the blood fluke *Schistosoma mansoni*. *Nature* **460**, 352–358 (2009).

30. Chen, L., Wilson, D. J., Labello, N. P., Jayaram, H. N. & Pankiewicz, K. W. Mycophenolic acid analogs with a modified metabolic profile. *Bioorg. Med. Chem.* **16**, 9340–9345 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This project was funded by the Australian Research Council. This research was supported by a Victorian Life Sciences Computation Initiative (grant number VR0007) on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government. Other support from the Australian Academy of Science, the Australian-American Fulbright Commission, Melbourne Water Corporation, and the IBM Research Collaboratory for Life Sciences—Melbourne is gratefully acknowledged. P.W.S. is an investigator with the Howard Hughes Medical Institute. A.R.J. held a CDA1 (Industry) from the National Health and Medical Research Council of Australia. We are indebted to the faculty and staff of the BGI-Shenzhen, who contributed to this study. We also acknowledge the contributions of staff at WormBase ([www.wormbase.org](http://www.wormbase.org)).

**Author Contributions** R.B.G., N.D.Y., B.E.C., J.V., T.W. and P.G. provided the samples and purified nucleic acids for sequencing. X.X. performed the whole genomic amplification of genomic DNA for the large insert libraries. S.L., L.Y., N.Z., A.R.J., Z.X., R.S.H., Y.K. and

F.C. undertook the sequencing, assembly and annotation of genomic and transcriptomic data. A.R.J., B.L., Z.X., N.D.Y., Y.L., R.S.H., E.M.S., G.Z., X.F., S.L., F.C. and C.C. planned and performed additional bioinformatic analyses. A.R.J., B.L., N.D.Y. and G.A.A. assisted with statistical analyses. A.R.J., P.N., E.M.S., P.W.S. and R.B.G. drafted and edited the manuscript, tables, figures and Supplementary Information. A.R.J., N.D.Y., S.R., J.W. and R.B.G. conceived and planned the project. A.R.J., B.L., N.D.Y., G.Z., X.F., X.W., J.W., Y.L., H.Y., J.W. and R.B.G. supervised and/or coordinated the research. T.W.H. curated the browsable genome.

**Author Information** All of the genomic sequence data have been released for public access at WormBase ([www.wormbase.org](http://www.wormbase.org)) and are accessible via [ftp://ftp.wormbase.org/pub/wormbase/species/a\\_suum](ftp://ftp.wormbase.org/pub/wormbase/species/a_suum). A browsable genome is also available (via [http://www.wormbase.org/db/gb2/gbrowse/a\\_suum/](http://www.wormbase.org/db/gb2/gbrowse/a_suum/)). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at [www.nature.com/nature](http://www.nature.com/nature). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to A.R.J. ([ajex@unimelb.edu.au](mailto:ajex@unimelb.edu.au)), J.W. ([wangjun@genomics.org.cn](mailto:wangjun@genomics.org.cn)), J.W. ([wangjian@genomics.org.cn](mailto:wangjian@genomics.org.cn)) or R.B.G. ([robinbg@unimelb.edu.au](mailto:robinbg@unimelb.edu.au)).

## METHODS

**Sample procurement, preparation and storage.** All specimens of *A. suum* were collected from pigs (*Sus scrofa*) with naturally acquired infections in Victoria, Australia (adult nematodes) and Ghent, Belgium (larval stages). L3s and L4s were also collected from the liver or lung and from the small intestine, respectively, of pigs, using established procedures<sup>31,32</sup>. Nematodes were washed extensively in sterile physiological saline (37 °C), snap-frozen in liquid nitrogen and then stored at -70 °C until use.

**DNA isolation, sequencing and quality control.** Total genomic DNA was isolated from the reproductive tract of a single adult female of *A. suum* using a sodium-dodecyl sulphate/proteinase K digestion<sup>33</sup> followed by phenol-chloroform extraction and ethanol precipitation<sup>34</sup>. Total DNA yield was determined using the Qubit fluorometer double-stranded DNA HS Kit (Invitrogen). DNA integrity was verified with a 2100 Bioanalyser (Agilent). Short-insert (170 bp and 500 bp) and mate-pair (800 bp, 2 kb, 5 kb and 10 kb) genomic DNA libraries were prepared and paired-end sequenced using TruSeq chemistry on a HiSeq 2000 (Illumina). Whole-genome amplification, employing the REPLI-g Midi Kit (Qiagen), was used to produce (from 200 ng of genomic template) the required amount of DNA for the construction of the 2-kb, 5-kb and 10-kb libraries (Supplementary Fig. 6). The sequence data generated from each of the six libraries were verified, and low-quality sequences, base-calling duplicates and adapters removed. The size of the genome and the heterozygosity rate were estimated by establishing the frequency of occurrence of each 17-bp *k*-mer (a unique sequence of *k* (that is, 17) nucleotides in length) within the genomic sequence data set (from the 170-bp library) using an established method<sup>5</sup>. Genome size was estimated using a modification of the Lander–Waterman algorithm<sup>35</sup>, where the haploid genome length in base pairs is  $G = (N \times (L - K + 1) - B)/D$ , where *N* is the read length sequenced in base pairs, *L* is the mean length of sequence reads, *K* is the *k*-mer length (17 bp) and *B* is the number of *k*-mers occurring less than four times (Supplementary Fig. 7). Heterozygosity was evaluated throughout the genome assembly by assessing the distribution of the *k*-mer frequency in the sequence data set.

**RNA isolation, sequencing and assembly.** We obtained total RNAs from egg-L3s (*n* ≈ 500,000), liver-L3s (*n* ≈ 60,000), lung-L3s (*n* ≈ 80,000) or L4s (*n* ≈ 30,000) and from the somatic musculature or reproductive tract of each of two adult male and two adult female *A. suum* using the TriPure reagent (Roche), and both yield and quality were verified by 2100 BioAnalyser (Agilent). Polyadenylated (polyA+) RNA was purified from 10 µg of total RNA using Sera-mag oligo(dT) beads, fragmented to a size of 300–500 bp, reverse-transcribed using random hexamers, end-repaired and adaptor-ligated, according to the manufacturer's protocol (Illumina). Ligated products of ~400 bp were excised from agarose and then PCR-amplified (15 cycles), as recommended. Products were purified over a MinElute column (Qiagen) and subjected to paired-end RNA-seq using TruSeq chemistry on a HiSeq 2000 (Illumina) and assessed for quality and adaptor sequence. Transcripts were assembled from RNA-seq data using Oases<sup>36</sup>. All transcripts were used to assess the completeness of the genome assembly and to predict genes.

**Genomic assembly and quality control.** Following sequencing, all DNA-sequence reads were corrected based on *k*-mer (=17) distribution<sup>5</sup>. Briefly, sequence reads were removed if >10% of bases were ambiguous (represented by the letter N) or multiple adenosine monophosphates (poly-A), and all remaining reads were filtered on the basis of Phred quality. For small insert-size libraries (that is, <800 bp), additional reads were removed from the final data set if >65% of bases were of a low Phred quality (<8). For large insert libraries (2 kb, 5 kb and 10 kb), reads were removed from the final data set if >80% of bases were of a low Phred quality (<8). Duplicate (that is, identical) reads and partial reads representing the Illumina adaptor sequence were also removed, as were reads from the 500-bp library representing paired reads found to overlap by >10 bp (allowing for a 10% mismatch). Corrected and filtered data were assembled into contigs using SOAPdenovo<sup>3</sup>, and joined iteratively into scaffolds using a step-wise process (see Supplementary Fig. 8), using the paired reads generated from each library; local assemblies were used to close all gaps. Each nucleotide position in the final assembly was assessed for accuracy by aligning all filtered reads to the scaffolds using SOAP2aligner<sup>37</sup>, allowing for up to five mismatches per read. The depth of coverage and repeat content were assessed initially by sliding-window analysis and presented as a frequency distribution (Supplementary Fig. 9). GC-content was estimated using 10-kb non-overlapping sliding windows, and GC-bias<sup>38</sup> was assessed based on a frequency distribution of these data (Supplementary Fig. 10). To assess the completeness of the genome assembly, RNA-seq data representing each of the organs (that is, musculature and reproductive tract), genders and/or stages of *A. suum* sequenced were mapped to the final assembly using the BLAST-like Alignment Tool (BLAT)<sup>39</sup>.

**Assessment of repeat content and annotation of non-coding RNA.** Following genome assembly, tandem repeats were identified using the Tandem Repeats Finder program<sup>40</sup>. Transposable elements were predicted using a combination of homology-based comparisons (using RepeatMasker<sup>41</sup>) and *de novo* approaches (using LTR\_FINDER<sup>42</sup>, PILER<sup>43</sup> and RepeatScout<sup>44</sup>), with a consensus population of predicted repetitive elements, constructed in RepeatScout using fit-preferred alignment scores. Low-frequency repeats (≤25) and multi-copy genes (in the repeat element library) were filtered using RepeatMasker, producing a non-redundant sequence file, which was then used to identify and classify additional homologous repeats in the genome.

**Gene prediction, and synteny and genetic variation analysis.** The *A. suum* protein-coding gene set was inferred using *de novo*-, homology- and evidence-based (that is, transcriptomic) approaches. *De novo* gene prediction was performed on a repeat-masked genome using three programs (Augustus, GlimmerHMM and SNAP)<sup>5</sup>; training models were generated from a subset of the transcriptomic data set representing 1,355 distinct genes. Homology-based prediction was conducted by comparison with complete genomic data for *Caenorhabditis elegans*<sup>3</sup>, *Pristionchus pacificus*<sup>8</sup> and *Brugia malayi*<sup>9</sup> using a multi-phase strategy, in which (1) all putative homologous gene sequences were preliminarily identified from alignments with protein sequences representing the complete gene set of each of the reference genomes (the longest transcripts were chosen to represent each gene) by TblastN (*e*-value cut-off: 10<sup>-5</sup>) and grouped into gene-like structures using genBlastA<sup>45</sup>; (2) regions representing these putative genes, and flanking regions (3,000 bp) at the 5'- and 3'-ends of each predicted gene, were extracted from the assembly and aligned to the 'parent' sequences derived from the reference genomes using Genewise<sup>46</sup>; (3) all single-exon genes predicted to have arisen from a retro-transposition and containing at least one frame-shift error or representing incomplete coding domains of <150 bp as well as all multi-exon genes containing more than two frame-shift errors and/or representing incomplete coding domains of <100 bp, were discarded. Evidence-based gene prediction was conducted by aligning all RNA-seq data generated herein against the assembled genome using TopHat<sup>47</sup>, with cDNAs predicted from the resultant data using Cufflinks<sup>48</sup>. Following the prediction of genes, a non-redundant gene set representing homology-based, *de novo*-predicted and RNA-seq-supported genes, was generated using Glean (<http://sourceforge.net/projects/glean-gene/>)<sup>5</sup>. All Glean-predicted genes were retained, as were all genes supported by RNA-seq data and those predicted using two or more *de novo* methods (that is, Augustus, GlimmerHMM and/or SNAP). The open reading frame of each gene was predicted using BestORF ([www.softberry.com](http://www.softberry.com)). To assess the quality and accuracy of the predicted gene set, we examined the length-distribution of all genes, coding sequences, exons and introns, and the distribution of exon numbers for individual genes, and then compared these parameters with those calculated for the published gene sets of *B. malayi*, *C. elegans*, *P. pristionchus* and *M. incognita* (Supplementary Fig. 4).

Following prediction of the finalized gene set, we conducted pairwise analysis of the overall synteny existing between/among the large (>1 Mb) assembly scaffolds for *B. malayi* and *A. suum* relative to the complete *C. elegans* chromosomes. This analysis was undertaken by conducting pairwise alignments among all *A. suum* or *B. malayi* (WS220 assembly: [ftp://ftp.sanger.ac.uk/pub2/wormbase/releases/WS220/genomes/b\\_malayi/](ftp://ftp.sanger.ac.uk/pub2/wormbase/releases/WS220/genomes/b_malayi/)) scaffolds larger than 1 Mb in size and the *C. elegans* chromosomes using LASTz ([http://www.bx.psu.edu/miller\\_lab/dist/README.lastz-1.02.00/README.lastz-1.02.00a.html](http://www.bx.psu.edu/miller_lab/dist/README.lastz-1.02.00/README.lastz-1.02.00a.html)), which were then joined using CHAINNET<sup>49</sup> and output as a .axt alignment from which large-synteny regions were defined. The resulting alignment files were used to construct synteny images on scaleable vector graphics format using customized perl scripts (ZX). In addition, gene-level synteny analyses were conducted for one-to-one orthologous genes colocalizing to large *A. suum* or *B. malayi* assembly scaffolds (>1 Mb) according to ref. 9. Orthology was determined by pairwise reciprocal BLASTx comparisons between *A. suum*, *B. malayi* and *C. elegans* according to ref. 11. One-to-one orthologous genes shared between either *A. suum* or *B. malayi* and *C. elegans* but not shared among *A. suum* and *B. malayi* based on reciprocal BLASTp analysis were further confirmed by Hidden Markov Modelling using the *jackhmmer* command in the program HMMER 3.0 (ref. 50) and a highly permissive threshold (HMM cutoff: 10<sup>-2</sup>).

We assessed the genome-wide variation in the exonic regions by mapping all raw reads from our transcriptomic data to the genomic coding domains using Maq<sup>51</sup>, and calling SNPs with a minimum coverage threshold of ten reads. All mapped reads were assessed as synonymous (non-coding change), non-synonymous (coding change) or ambiguous (a SNP that was represented in our data set as an ambiguous IUPAC code wherein one nucleotide change would cause a synonymous mutation and the other a non-synonymous mutation) using a custom Perl script (snp\_analysis.pl). All genes were then ranked based on their accumulation of SNPs to assess and identify their levels of conservation/variation relative to their

function. We reasoned that, in addition to the real effects of the variability of each gene on their accumulation of SNPs, these data would be influenced also by the coverage achieved for each gene, which is affected by the number of reads available for each gene (that is, their relative levels of transcription) and the length of each gene. Thus, before ranking, the SNP data for each gene was normalized for its calculated reads per kilobase per million reads (RPKM) and total gene length using the simple equation: SNPs per read per kilobase = total SNPs divided by RPKM divided by gene length (in bp) multiplied by 1,000 bp. Following ranking, we explored function among the 2.5% most variable (with the highest rankings based on normalized SNP data) and most conserved genes (with the lowest rankings based on normalized SNP data). Noting the potential inaccuracy associated with estimating the normalized SNP rankings of lowly transcribed genes (owing to a lack of data/coverage), only genes for which at least 100 reads were available were considered in these functional comparisons.

**Functional annotation of coding genes.** Following the prediction of the protein-coding gene set, each inferred amino acid sequence was assessed for conserved protein domains in the SPOT, Pfam, PRINTS, PROSITE, ProDom and SMART databases using InterProScan<sup>52</sup>, employing default settings. Gene ontology categories<sup>53</sup> were assigned to each contig inferred to contain at least one conserved protein domain. Gene ontology categories were summarized and standardized to level 2 and level 3 terms, defined using the GOSlim hierarchy<sup>54</sup> using WEGO<sup>55</sup>. To characterize further the contigs/transcripts from *A. suum*, we conducted a series of high-stringency BLASTp homology searches (*e*-value cut-off:  $10^{-5}$ ) against a variety of databases. Each contig was assessed for a known functional orthologue, defined using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (www.kegg.com). Where appropriate, orthologous matches were mapped visually to a defined pathway using the KEGG pathway tool (available via www.kegg.com) or clustered to a known protein family using the KEGG-BRITE hierarchy tool (available via www.kegg.com). In addition, the amino acid sequence inferred from each *A. suum* coding gene was compared by BLASTp with protein sequences available for key nematode species (*B. malayi*, *C. elegans*, *P. pacificus* and *M. incognita*) as well as for *Drosophila melanogaster*<sup>4</sup> and *Mus musculus*<sup>56</sup> and those contained within the UniProt<sup>57</sup>, SwissProt and TrEMBL databases<sup>58</sup>. Key protein groups (for example, peptidases, kinases, phosphatases, GTPases, GPCRs, and transport and channel proteins) were characterized by high-stringency BLASTp homology searching (*e*-value cut-off  $<10^{-5}$ ) of manually curated information sequence data available in the MEROPS<sup>59</sup>, WormBase, KS-Sarfari (https://www.ebi.ac.uk/chembl/sarfari/kinasesarfari) and GPCR-Sarfari (https://www.ebi.ac.uk/chembl/sarfari/gpcrsarfari) and the Transporter Classification database<sup>60</sup>. E/S proteins were predicted using Phobius<sup>61</sup>, employing both the neural network and hidden Markov models, and by BLASTp homology-searching of the validated signal peptide database<sup>62</sup> and an E/S database containing published proteomic data for *B. malayi*<sup>63,64</sup>, *Schistosoma mansoni*<sup>65</sup> and *M. incognita*<sup>66</sup>. In the final annotation, proteins inferred from genes were classified based on a homology match (*e*-value cut-off:  $\leq 10^{-5}$ ) to: (1) a curated, specialist protein database, followed by (2) the KEGG database, followed by (3) the UniProt/SwissProt/TrEMBL databases, followed by (4) the annotated gene set for a model organism, including *C. elegans*, *D. melanogaster*, *M. musculus* or *S. cerevisiae*, followed by (5) the gene ontology classification, and, finally, (6) a recognized, conserved protein domain based on InterProScan analysis. Any inferred proteins lacking a match (BLASTp cut-off  $\leq 10^{-5}$ ) in at least one of these analyses were designated hypothetical proteins. The final annotated protein-coding gene set for *A. suum* is available for download at WormBase (in nucleotide and amino acid formats).

**Differential transcription analysis.** Following RNA-seq, all paired-end reads for each library constructed were aligned to the predicted *A. suum* gene set using TopHat, and quantitative levels of transcription (RPKM)<sup>67</sup> were calculated using Cufflinks. Differential transcription was assessed<sup>68</sup> using a *P*-value cut-off of  $\leq 0.01$  and a minimum, two-fold difference in absolute RPKM values. False discovery rates for differential transcription were determined<sup>68</sup>. To allow the rapid visual assessment of the statistically significant changes in transcription of each gene between and among individual libraries, we constructed heat-maps representing absolute differences in the RPKM values, calculated for each transcript using a customized Perl script (express\_heatmap\_RPKM.pl). Genetic interaction networks were predicted<sup>69</sup> based on data available for homologous genes in *C. elegans* (inferred from BLASTp comparisons) and viewed using the program BioLayout 3D<sup>70</sup>.

**Essentiality and druggability predictions.** *A. suum* genes with homology to those in the *C. elegans* and/or *D. melanogaster* genomes were inferred based on BLASTp comparisons using the predicted protein sequences for individual species (*e*-value cut-off  $10^{-5}$ ). Phenotypic data for each *C. elegans* and *D. melanogaster* homologue were sourced from WormBase and FlyBase (www.flybase.org), respectively. *A. suum* genes determined<sup>71</sup> to have homologues with lethal phenotypes in both *C. elegans* and *D. melanogaster* were inferred to represent essential

genes. Metabolic chokepoints were defined<sup>29,72</sup> and assessed based on *A. suum* gene sequences determined, by BLASTp comparison ( $10^{-5}$ ), to have an orthologue in the KEGG database. All 'essential' homologues and/or molecules in 'choke-points' were then queried against the BRENDA<sup>73</sup> and ChEMBL databases (accessible via https://www.ebi.ac.uk/chembl/bd/), to identify known chemical inhibitors.

**Additional bioinformatic analyses, and use of software.** Data analysis was conducted in a Unix environment or Microsoft Excel 2007 using standard commands. Bioinformatic scripts required to facilitate data analysis were designed using Perl, BioPerl, Java and Python and are available via http://research.vet.unimelb.edu.au/gasserlab/.

- Cantacessi, C. et al. Differences in transcription between free-living and CO<sub>2</sub>-activated third-stage larvae of *Haemonchus contortus*. *BMC Genom.* **11**, 266, doi:10.1186/1471-2164-11-266 (2010).
- Saeed, I., Roepstorff, A., Rasmussen, T., Hog, M. & Jungersen, G. Optimization of the agar-gel method for isolation of migrating *Ascaris suum* larvae from the liver and lungs of pigs. *Acta Vet. Scand.* **42**, 279–286 (2001).
- Gasser, R. B. et al. Single-strand conformation polymorphism (SSCP) for the analysis of genetic variation. *Nature Protocols* **1**, 3121–3128 (2007).
- Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual* 3rd edn, Vol. 3, E.3–E.4 (Cold Spring Harbor Laboratory, 2001).
- Lander, E. S. & Waterman, M. S. Genomic mapping by fingerprinting random clones: a mathematical analysis. *Genomics* **2**, 231–239 (1988).
- Zerbino, D. R. & Birney, E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
- Li, R. et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967 (2009).
- Bentley, D. R. et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
- Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
- Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
- Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protocols Bioinformatics* Ch. 4.10 (2009).
- Xu, Z. & Wang, H. LTR\_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).
- Edgar, R. C. & Myers, E. W. PILER: identification and classification of genomic repeats. *Bioinformatics* **21** (Suppl. 1), i152–i158 (2005).
- Price, A. L., Jones, N. C. & Pevzner, P. A. *De novo* identification of repeat families in large genomes. *Bioinformatics* **21** (Suppl. 1), i351–i358 (2005).
- She, R., Chu, J. S., Wang, K., Pei, J. & Chen, N. GenBlastA: enabling BLAST to identify homologous gene sequences. *Genome Res.* **19**, 143–149 (2009).
- Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).
- Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
- Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnol.* **28**, 511–515 (2010).
- Kent, W. J., Baertsch, R., Hinrichs, A., Miller, W. & Haussler, D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proc. Natl Acad. Sci. USA* **100**, 11484–11489 (2003).
- Eddy, S. R. A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **23**, 205–211 (2009).
- Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18**, 1851–1858 (2008).
- Quevillon, E. et al. InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, W116–W120 (2005).
- Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, D258–D261 (2004).
- Camon, E. et al. The Gene Ontology Annotation (GOA) project: implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Res.* **13**, 662–672 (2003).
- Ye, J. et al. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* **34**, W293–W297 (2006).
- Chinwalla, A. T. et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
- Wu, C. H. et al. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* **34**, D187–D191 (2006).
- Boeckmann, B. et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**, 365–370 (2003).
- Rawlings, N. D., Barrett, A. J. & Bateman, A. MEROPS: the peptidase database. *Nucleic Acids Res.* **38**, D227–D233 (2010).
- Saier, M. H. Jr, Yen, M. R., Noto, K., Tamang, D. G. & Elkan, C. The Transporter Classification Database: recent advances. *Nucleic Acids Res.* **37**, D274–D278 (2009).
- Kall, L., Krogh, A. & Sonnhammer, E. L. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.* **35**, W429–W432 (2007).
- Chen, Y. et al. SPD—a web-based secreted protein database. *Nucleic Acids Res.* **33**, D169–D173 (2005).
- Bennuru, S. et al. *Brugia malayi* excreted/secreted proteins at the host/parasite interface: stage- and gender-specific proteomic profiling. *PLoS Negl. Trop. Dis.* **3**, e410 (2009).

64. Hewitson, J. P. *et al.* The secretome of the filarial parasite, *Brugia malayi*: proteomic profile of adult excretory-secretory products. *Mol. Biochem. Parasitol.* **160**, 8–21 (2008).
65. Cass, C. L. *et al.* Proteomic analysis of *Schistosoma mansoni* egg secretions. *Mol. Biochem. Parasitol.* **155**, 84–93 (2007).
66. Bellafiore, S. *et al.* Direct identification of the *Meloidogyne incognita* secretome reveals proteins with host cell reprogramming potential. *PLoS Pathog.* **4**, e1000192 (2008).
67. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).
68. Audic, S. & Claverie, J. M. The significance of digital gene expression profiles. *Genome Res.* **7**, 986–995 (1997).
69. Zhong, W. & Sternberg, P. W. Genome-wide prediction of *C. elegans* genetic interactions. *Science* **311**, 1481–1484 (2006).
70. Goldovsky, L., Cases, I., Enright, A. J. & Ouzounis, C. A. BioLayout(Java): versatile network visualisation of structural and functional relationships. *Appl. Bioinform.* **4**, 71–74 (2005).
71. Doyle, M. A., Gasser, R. B., Woodcroft, B. J., Hall, R. S. & Ralph, S. A. Drug target prediction and prioritization: using orthology to predict essentiality in parasite genomes. *BMC Genom.* **11**, 222, doi:10.1186/1471-2164-11-222 (2010).
72. Yeh, I., Hanekamp, T., Tsoka, S., Karp, P. D. & Altman, R. B. Computational analysis of *Plasmodium falciparum* metabolism: organizing genomic information to facilitate drug discovery. *Genome Res.* **14**, 917–924 (2004).
73. Scheer, M. *et al.* BRENDA, the enzyme information system in 2011. *Nucleic Acids Res.* **39**, D670–D676 (2011).

# Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination

Kerstin Berer<sup>1</sup>, Marsilius Mues<sup>1</sup>, Michail Koutouros<sup>1</sup>, Zakeya Al Rasbi<sup>1</sup>, Marina Boziki<sup>1</sup>, Caroline Johner<sup>2</sup>, Hartmut Wekerle<sup>1</sup> & Gurumoorthy Krishnamoorthy<sup>1</sup>

Active multiple sclerosis lesions show inflammatory changes suggestive of a combined attack by autoreactive T and B lymphocytes against brain white matter<sup>1</sup>. These pathogenic immune cells derive from progenitors that are normal, innocuous components of the healthy immune repertoire but become autoaggressive upon pathological activation. The stimuli triggering this autoimmune conversion have been commonly attributed to environmental factors, in particular microbial infection<sup>2</sup>. However, using the relapsing–remitting mouse model of spontaneously developing experimental autoimmune encephalomyelitis<sup>3</sup>, here we show that the commensal gut flora—in the absence of pathogenic agents—is essential in triggering immune processes, leading to a relapsing–remitting autoimmune disease driven by myelin-specific CD4<sup>+</sup> T cells. We show further that recruitment and activation of autoantibody-producing B cells from the endogenous immune repertoire depends on availability of the target autoantigen, myelin oligodendrocyte glycoprotein (MOG), and commensal microbiota. Our observations identify a sequence of events triggering organ-specific autoimmune disease and these processes may offer novel therapeutic targets.

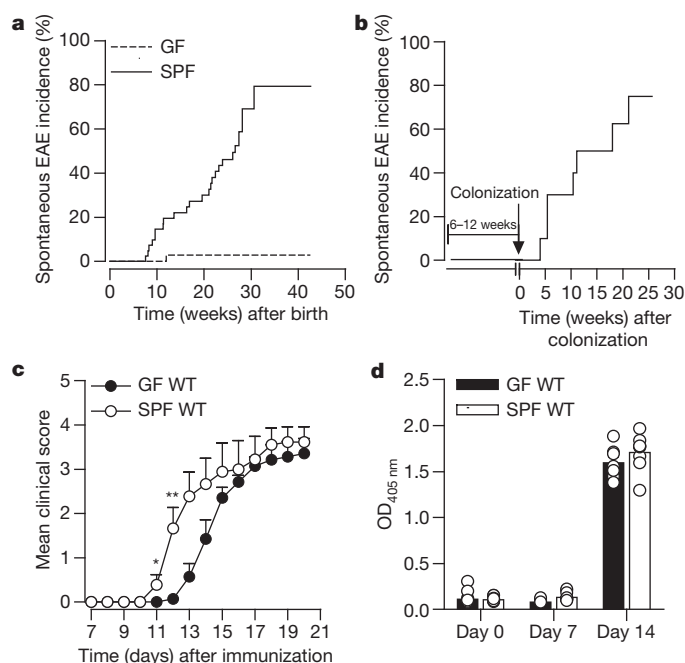
The relapsing–remitting (RR) mouse model uses transgenic SJL/J mice expressing, in a large proportion of their CD4<sup>+</sup> T cells, a transgenic T-cell antigen receptor (TCR) recognizing MOG peptide 92–106 in the context of MHC class II, I-A<sup>s</sup>. These mice spontaneously develop experimental autoimmune encephalomyelitis (EAE) with successive disease bouts that often affect different central nervous system (CNS) tissues. The disease is initiated by the transgenic CD4<sup>+</sup> T cells, which first infiltrate the CNS, and by MOG-autoantibody-producing B cells recruited from the natural immune repertoire<sup>3</sup>.

Whereas in our facility close to 80% of RR mice developed spontaneous EAE within 3–8 months of age, the rate was variable in other institutions, with spontaneous EAE incidences ranging from 35–90% (unpublished data). This recalled previous investigations that also observed that the frequency of spontaneous EAE in myelin-specific TCR transgenic mice varied in different breeding centres<sup>4</sup>. Because our mice were reared under specific pathogen-free (SPF) conditions, we tested the possible contributions of the non-pathogenic commensal flora to the triggering of a spontaneous CNS-specific autoimmune disease.

We first compared the incidence of spontaneous EAE between RR mice housed under SPF and completely germ-free conditions. The differences were marked. Whereas, as reported before, most SPF-bred RR mice came down with EAE within 3–8 months<sup>3</sup>, germ-free RR mice remained fully protected throughout their life (Fig. 1a). As the commensal microbiota have a central function in driving the correct development of the immune system<sup>5</sup>, the absence of spontaneous EAE in germ-free RR mice may have reflected a general immune deficiency due to missing microbial stimuli. However, two observations argue against a profound and irreversible non-reactivity. First, RR mice, which had been germ free (and disease free) for 6–12 weeks, promptly developed EAE when re-colonized with conventional commensal microbiota (Fig. 1b). Mono-colonization with segmented filamentous

bacteria (SFB), which restored autoimmunity in another mouse model, was of low efficiency (unpublished data). This suggests that the immune system of germ-free mice had grown efficient enough to mount a full autoimmune attack within a relatively brief period of time. Second, the basic immune competence of germ-free animals was confirmed by active immunization of germ-free wild-type SJL/J mice with recombinant MOG (rMOG) in complete Freund's adjuvant (CFA). In accord with one previous report<sup>6</sup>, although not with another more recent one<sup>7</sup>, all immunized germ-free mice developed EAE like their SPF counterparts, although with some delay (Fig. 1c), and transfer of pre-activated T cells induced comparable EAE in both germ-free and SPF mice (Supplementary Table 1). Moreover, germ-free and SPF SJL/J mice immunized with rMOG produced comparable levels of anti-MOG antibodies in their serum (Fig. 1d).

Recent studies established that components of the commensal microbiota profoundly shape the gut-associated lymphatic tissue (GALT),



**Figure 1 | Commensal microbiota are required for the development of spontaneous EAE.** **a**, Incidence of spontaneous EAE in a cohort of RR mice housed in germ-free (GF;  $n = 35$ ) or SPF ( $n = 41$ ) conditions. **b**, Incidence of spontaneous EAE in germ-free RR mice ( $n = 10$ ) re-colonized with conventional flora from SPF mice. **c**, Delayed EAE onset in germ-free wild-type (GF WT) SJL/J mice immunized with rMOG/CFA. Mean EAE scores ( $\pm$  s.e.m.) of germ-free ( $n = 7$ ) and SPF ( $n = 8$ ) SJL/J mice are shown. \* $P < 0.05$ ; \*\* $P < 0.01$  (two-way ANOVA). **d**, Germ-free and SPF wild-type SJL/J mice produce similar levels of anti-MOG antibodies after immunization. Each circle represents an individual mouse and bars depict mean  $\pm$  s.e.m. Panels **c** and **d** represent two individual experiments.

<sup>1</sup>Department of Neuroimmunology, Max Planck Institute of Neurobiology, 82152 Martinsried, Germany. <sup>2</sup>Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany.

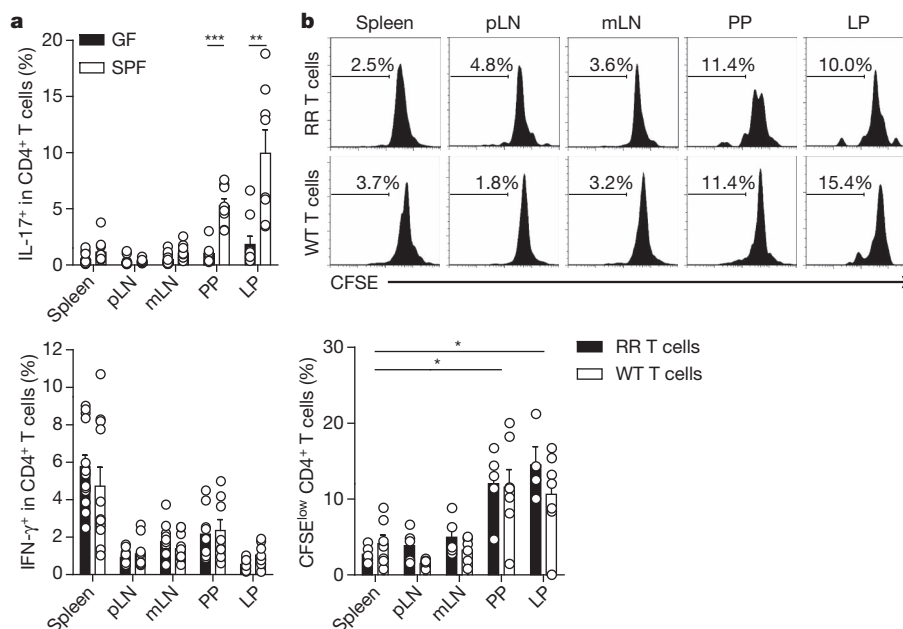
some supporting differentiation of interleukin (IL)-17-producing  $T_H17$  cells<sup>8–10</sup> and others the generation of regulatory T lymphocytes ( $T_{reg}$ )<sup>11,12</sup>. We also found a marked deficit of  $T_H17$ -like cells in germ-free mice (Fig. 2a), which was most pronounced in T cells intimately connected to the intestinal wall, lamina propria T cells and in Peyer's patch but not in mesenteric lymph node populations. There were no notable changes in remote organs, such as spleen or pooled inguinal and axillary lymph nodes (Fig. 2a). Frequencies of IFN- $\gamma$ -, TNF- $\alpha$ - and IL-10-producing  $CD4^+$  T cells were comparable between germ-free and SPF RR mice (Fig. 2a and Supplementary Fig. 1). Apart from a minor increase in the frequency of  $CD4^+$  T cells in the spleen of germ-free RR mice and a reduction of the T cells expressing lower levels of T-cell receptor ( $TCR\alpha\beta^{low}$ ) (activated T cells) in the lamina propria (Supplementary Figs 2 and 3), the proportions of most lymphoid cell types examined, including  $Foxp3^+$   $T_{reg}$  cells,  $CD8^+$  T cells,  $TCR\gamma\delta^+$  cells, B cells,  $CD11b^+$  macrophages,  $CD11c^+$  dendritic cells, natural killer (NK) cells and  $Gr1^+$  granulocytes were unchanged (Supplementary Fig. 2). Of note, although in the spleen the microbial colonization status did not affect cellular composition, it definitely impinged on cytokine production of splenic immune cells. As in MOG-immunized germ-free C57BL/6 mice<sup>7</sup>, germ-free RR mouse spleen cells released lower levels of IL-17 than their SPF counterparts upon MOG antigen or anti-CD3 monoclonal antibody stimulation, and in addition they showed reduced secretion of IFN- $\gamma$ . Re-colonization of germ-free mice not only restored T-cell cytokine production capacity but even led to overshooting reactions (Supplementary Fig. 4).

The commensal microbiota could act on MOG-specific T cells either via microbial structures mimicking MOG epitopes<sup>13</sup> or through innate immune signals creating a particular inflammatory milieu<sup>11,14</sup>. In an attempt to probe a potential MOG-specific mimicry response, we transferred carboxyfluorescein succinimidyl ester (CFSE)-labelled TCR transgenic or wild-type T cells into SPF wild-type mice and tested their proliferative responses in the gut. Proliferation rates of transgenic and polyspecific wild-type T cells in the GALT were equally high, whereas in the remote spleen of the same recipients the responses remained hardly detectable (Fig. 2b). Further, the microbial signals

seem to act persistently on local T cells. Transient depletion of gut flora by short-term antibiotic treatment significantly reduced the proliferation of T cells in the lamina propria, but not in spleen, pooled lymph nodes, mesenteric lymph nodes and Peyer's patches (Supplementary Fig. 5).

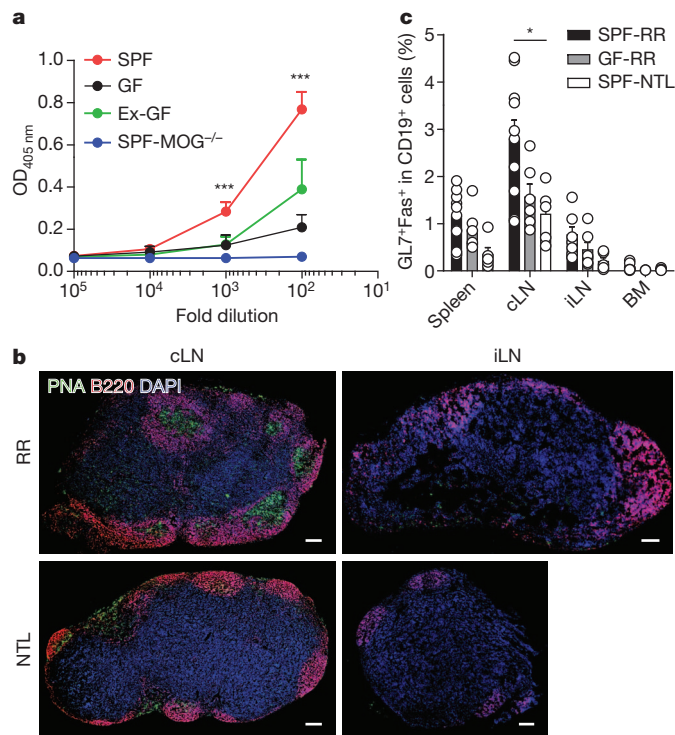
Activation of MOG-specific T cells in the GALT is necessary for the development of EAE in RR mice, but not sufficient. Full clinical EAE requires the participation of MOG-reactive B lymphocytes. We proposed that in RR mice, transgenic pathogenic T cells select the auto-immune B cells from the native B-cell repertoire and drive them to proliferate and release autoantibodies of IgG classes<sup>3</sup>. Indeed, we now found that germ-free RR mice, which, owing to missing microbial stimuli, lack activated autoimmune T cells, produced only low doses of anti-MOG autoantibodies. The autoantibody production was promptly increased in germ-free mice upon re-colonization (Fig. 3a). This response could involve antigenic mimicry at the B-cell level between MOG and epitopes on commensal microbes, reminiscent of Sydenham's chorea—the CNS manifestation of rheumatic fever—in which streptococcal antigens mimic neuronal B-cell epitopes<sup>15</sup>. However, this was not the case in spontaneous RR mouse EAE. We discovered that production of demyelinating autoantibodies critically depended on the expression of the target myelin autoantigen, MOG. RR mice deficient in MOG ( $RR \times MOG^{-/-}$ ), due to a transgenic mutation of the *Mog* gene<sup>16</sup>, failed to develop anti-MOG autoantibody titres despite their normal microbial status (Fig. 3a). Importantly, our data show that exogenous MOG injected into SPF  $RR \times MOG^{-/-}$  mice via MOG in CFA readily induced anti-MOG antibodies (Supplementary Fig. 6).

Recruitment and activation of antigen-specific B cells involves signals by local helper T cells and surrounding stroma cells, which together drive the resting B cell into a germinal centre, where it undergoes proliferation, immunoglobulin class switching and somatic hypermutation<sup>17</sup>. Binding of the cognate antigen to the B-cell receptor has a central role in these processes. MOG-specific B cells could be recruited either in the CNS tissue via locally produced MOG material<sup>18</sup>, or in CNS draining peripheral lymph nodes (deep cervical lymph



**Figure 2 | Effect of microbiota on T-cell activation and their cytokine profiles in the GALT. a**, Impaired  $T_H17$  differentiation in germ-free (GF) RR mice. Frequencies of IL-17- or IFN- $\gamma$ -producing T cells from the indicated organs of GF and SPF RR mice are shown. LP, lamina propria; mLN, mesenteric lymph nodes; pLN, pooled inguinal and axillary lymph nodes; PP, Peyer's patches.  $n = 8$ –13 mice per group. Data were pooled from four independent

experiments. \*\*\* $P = 0.0002$ ; \*\* $P = 0.0025$  (Mann–Whitney U test). **b**, Activation of T cells by commensal flora. Shown are the frequencies of CFSE<sup>low</sup>  $CD4^+$  cells in the indicated organs of mice that received CFSE-labelled  $CD4^+$  T cells. Each circle represents an individual mouse and bars depict mean  $\pm$  s.e.m.  $n = 4$ –7 mice per group. Data represent two individual experiments. \* $P < 0.05$  (Mann–Whitney U test).

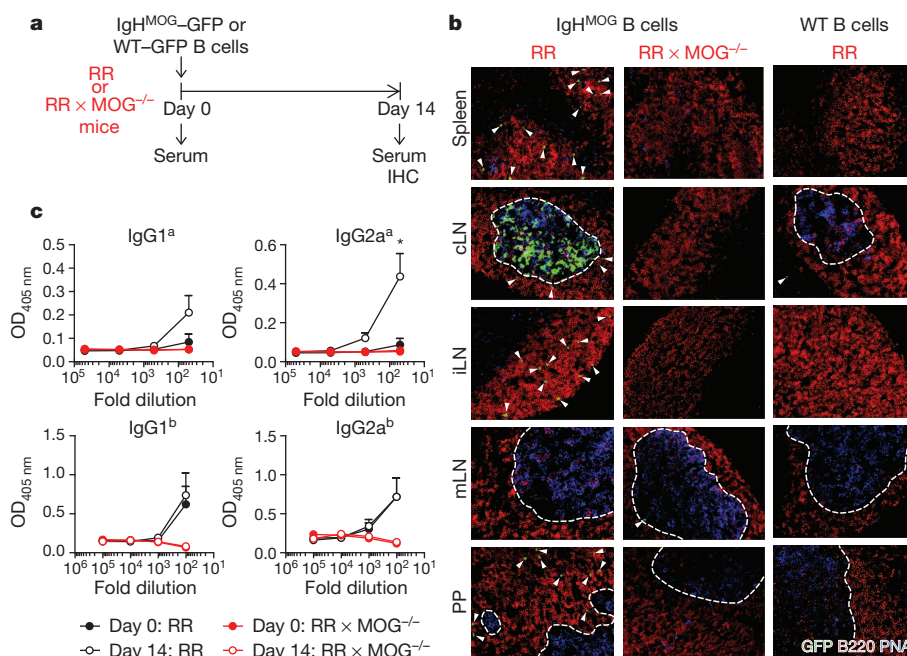


nodes) with MOG imported from the CNS via lymphatic vessels<sup>19</sup>. Our observations favour the latter alternative. Prior to the onset of clinical symptoms in SPF RR mice we found some scattered B cells in CNS infiltrates (Supplementary Fig. 7) but no follicle-like aggregates or follicular markers (data not shown). However, there were conspicuous changes in the cervical lymph nodes. Starting from the age of 3 weeks, cervical lymph nodes contained clearly delineated germinal centres (Fig. 3b). Furthermore, germinal centres and increased frequencies

**Figure 3 | B-cell recruitment is impaired in germ-free RR mice.** **a**, Reduced production of MOG-specific IgG2a antibodies in sera of germ-free (GF) RR mice. SPF ( $n = 15$ ); germ-free ( $n = 24$ ); germ-free mice re-colonized with conventional commensal microbiota (Ex-GF;  $n = 7$ ); or SPF-MOG<sup>-/-</sup> ( $n = 13$ ) RR mice. Error bars indicate s.e.m. \*\*\* $P < 0.001$  (Kruskal–Wallis test). **b**, **c**, Spontaneously formed germinal centre B cells are enriched in the cervical lymph nodes (cLN) of RR mice. **b**, Immunofluorescence staining of cervical lymph nodes or inguinal lymph node (iLN) sections of RR and NTL mice. Peanut agglutinin (PNA) (green; germinal centre), B220 (red; B cells) and 4',6-diamidino-2-phenylindole (DAPI) (blue; cell nuclei). Representative data of 4–5 individual mice are shown. Scale bars, 100  $\mu$ m. **c**, Flow cytometric analysis of GL7<sup>+</sup> Fas<sup>+</sup> germinal centre B cells from spleen, cervical lymph nodes, inguinal lymph nodes and bone marrow (BM). Each circle represents an individual mouse and bars represent mean  $\pm$  s.e.m.  $n = 5$ –10 per group. Data were pooled from three independent experiments. \* $P < 0.05$  (Mann–Whitney U test).

of GL7<sup>+</sup> and Fas<sup>+</sup> germinal centre B cells were restricted to cervical lymph nodes of RR mice, but were significantly reduced in age-matched germ-free and non-transgenic littermates (NTL) (Fig. 3b, c).

Germinal centres are attractive milieus for B cells, provided they contain appropriate antigenic material and competent T-helper cells<sup>20</sup>. To investigate whether RR mouse cervical lymph nodes offer both prerequisites to MOG-reactive B cells, we transferred GFP-labelled, MOG-reactive B cells expressing the H chain of a MOG monoclonal antibody (IgH<sup>MOG</sup>)<sup>21</sup> into hosts with a distinct allotype and traced their homing behaviour (Fig. 4a). When examined 14 days after transfer, MOG-specific B cells were found densely packed within germinal centres of cervical lymph nodes (Fig. 4b). Further, donor IgH<sup>MOG</sup> B cells switched anti-MOG antibodies to IgG2a isotypes (Fig. 4c). However, GFP-labelled polyclonal wild-type B cells failed to accumulate in any of the lymph nodes of RR mice (Fig. 4b), and IgH<sup>MOG</sup> B cells neither homed to the cervical lymph nodes of wild-type or RR  $\times$  MOG<sup>-/-</sup> mice (Fig. 4b and Supplementary Fig. 8) nor produced class-switched anti-MOG antibodies (Fig. 4c and Supplementary Fig. 8c). Collectively, these data indicate an ongoing MOG-specific germinal centre reaction, which is critically dependent on the expression of MOG, in the cervical lymph nodes of RR mice.



**Figure 4 | MOG-specific B cells home to the germinal centre of brain draining cervical lymph nodes.** **a**, Schematic representation of experimental set up. IHC, immunohistochemistry. **b**, MOG-specific B cells home to the germinal centre of brain draining cervical lymph nodes. PNA (blue; germinal centre) and B220 (red; B cells). Arrowheads indicate transferred GFP<sup>+</sup> B cells. Dotted lines define boundaries of germinal centre. Representative staining of

two independent experiments are shown.  $n = 4$ –5 mice per group. Magnification:  $\times 20$ . **c**, Transgenic MOG-specific B cells spontaneously switch isotype in RR but not in RR  $\times$  MOG<sup>-/-</sup> mice. Titres of donor (a allotype) and recipient (b allotype) anti-MOG antibodies were measured. Error bars represent s.e.m.  $n = 4$ –5 mice per group. Data were pooled from three independent experiments. \* $P < 0.05$  (Mann–Whitney U test).

This is the first report, to our knowledge, describing the sequential roles of the intact commensal gut flora and of myelin autoantigen in the initiation of a complex spontaneous demyelinating autoimmune disease. We propose a two-phase scenario that starts out in the GALT with expanding and activating CNS autoreactive T cells, which then recruit autoantibody-producing B cells. Together the autoimmune T and B cells trigger a demyelinating encephalomyelitis, which in the RR SJL/J mouse takes a relapsing–remitting course, very similar to early human multiple sclerosis.

Our findings are of direct relevance to multiple sclerosis, the pathogenesis of which is presently hotly debated. Some propose primary changes in the CNS target as the initiating process<sup>22</sup>, whereas others suggest that pathogenesis originates in the immune system<sup>23</sup>. Our present data support the latter concept. It is tempting to extend our finding of the gut origin of experimental CNS autoimmunity to human multiple sclerosis. There is now emerging evidence implicating gut microbiota in the starting phase of human autoimmune diseases. Besides inflammatory bowel diseases, in which bacteria may act on local tissue directly as well as indirectly<sup>24</sup>, inflammatory diseases with remote tissues affected seem to be modulated by the gut environment; for example, in rheumatoid arthritis<sup>25</sup> and type 1 diabetes mellitus<sup>26</sup>. In multiple sclerosis, evidence for commensal microbial contributions has remained less clear, so far. Dietary risk factors have been suggested to have a role<sup>27</sup>, and may contribute to a conspicuous increase of multiple sclerosis prevalence in Asian countries, like Japan, which has been ascribed to the spreading of a ‘westernized’ lifestyle<sup>28</sup>. It will be of interest now to search for the composition of intestinal microbiota associated with an increased susceptibility to multiple sclerosis, and this may provide a conceptual basis for exploring new, non-invasive treatment strategies.

## METHODS SUMMARY

**Mice.** Germ-free animals were re-derived from SJL/J anti-MOG TCR transgenic RR mice and kept germ free at the animal facility of the Max Planck Institute of Immunobiology and Epigenetics. Mice were re-colonized by housing in bedding material pre-conditioned by SPF mice.

**Cell purification, flow cytometry and adoptive transfers.** Single-cell suspensions were prepared from spleen, lymph nodes, Peyer’s patches and lamina propria by enzymatic digestion or mechanical disruption. Untouched T cells and B cells were purified using negative isolation kits (R&D Systems). Cells were stained with fluorochrome-labelled antibodies and acquired on FACSCalibur (BD Biosciences). Data were analysed using FlowJo (TreeStar) software. CFSE-labelled T cells or GFP<sup>+</sup> B cells were injected intravenously into SPF mice. CFSE<sup>low</sup> T cells were quantified by FACS 3 days after transfer. Localization of GFP<sup>+</sup> B cells was documented by immunofluorescence after 2 weeks.

**Immunofluorescence.** Sections of immune organs were stained with PNA (Vector Laboratories), anti-mouse B220 (BD Biosciences) and DAPI (Invitrogen). Images were obtained with a fluorescence microscope (Axiovert 200M; Carl Zeiss) and processed with MetaMorph 7.7 Software and Adobe Photoshop CS4.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 16 June; accepted 12 September 2011.

Published online 26 October 2011.

1. Franciotta, D. *et al.* B cells and multiple sclerosis. *Lancet Neurol.* **79**, 852–858 (2008).
2. Ascherio, A. & Munger, K. L. Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Ann. Neurol.* **61**, 288–299 (2007).
3. Pöhlner, B. *et al.* Spontaneous relapsing–remitting EAE in the SJL/J mouse: MOG-reactive transgenic T cells recruit endogenous MOG-specific B cells. *J. Exp. Med.* **206**, 1303–1316 (2009).
4. Goverman, J. *et al.* Transgenic mice that express a myelin basic protein-specific T cell receptor develop spontaneous autoimmunity. *Cell* **72**, 551–560 (1993).
5. Hill, D. A. & Artis, D. Intestinal bacteria and the regulation of immune cell homeostasis. *Annu. Rev. Immunol.* **28**, 623–667 (2010).

6. Lampropoulou, V. *et al.* TLR-activated B cells suppress T cell-mediated autoimmunity. *J. Immunol.* **180**, 4763–4773 (2008).
7. Lee, Y. K., Menezes, J. S., Umesaki, Y. & Mazmanian, S. K. Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. *Proc. Natl Acad. Sci. USA* **108**, 4615–4622 (2011).
8. Ivanov, I. I. *et al.* Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe* **4**, 337–349 (2008).
9. Gaboriau-Routhiau, V. *et al.* The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity* **31**, 677–689 (2009).
10. Ivanov, I. I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009).
11. Hall, J. A. *et al.* Commensal DNA limits regulatory T cell conversion and is a natural adjuvant of intestinal immune responses. *Immunity* **29**, 637–649 (2008).
12. Atarashi, K. *et al.* Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science* **331**, 337–341 (2011).
13. Ercolini, A. M. & Miller, S. D. Molecular mimics can induce novel self peptide-reactive CD4<sup>+</sup> T cell clonotypes in autoimmune disease. *J. Immunol.* **179**, 6604–6612 (2007).
14. Feng, T. *et al.* Microbiota innate stimulation is a prerequisite for T cell spontaneous proliferation and induction of experimental colitis. *J. Exp. Med.* **207**, 1321–1332 (2010).
15. Kirvan, C. A., Cox, C. J., Swedo, S. E. & Cunningham, M. W. Tubulin is a neuronal target of autoantibodies in Sydenham’s chorea. *J. Immunol.* **178**, 7412–7421 (2007).
16. Delarasse, C. *et al.* Myelin/oligodendrocyte glycoprotein-deficient (MOG-deficient) mice reveal lack of immune tolerance to MOG in wild-type mice. *J. Clin. Invest.* **112**, 544–553 (2003).
17. Cyster, J. G. B cell follicles and antigen encounters of the third kind. *Nature Immunol.* **11**, 989–996 (2010).
18. McMahon, E. J. *et al.* Epitope spreading initiates in the CNS in two mouse models of multiple sclerosis. *Nature Med.* **11**, 335–339 (2005).
19. Cserr, H. F., Harling-Berg, C. J. & Knopf, P. M. Drainage of brain extracellular fluid into blood and deep cervical lymph and its immunological significance. *Brain Pathol.* **2**, 269–276 (1992).
20. Schwickert, T. A., Alabyev, B., Manser, T. & Nussenzweig, M. C. Germinal center reutilization by newly activated B cells. *J. Exp. Med.* **206**, 2907–2914 (2009).
21. Litznerburger, T. *et al.* B lymphocytes producing demyelinating autoantibodies: development and function in gene-targeted transgenic mice. *J. Exp. Med.* **188**, 169–180 (1998).
22. Barnett, M. H., Henderson, A. P. D. & Prineas, J. W. The macrophage in MS: just a scavenger after all? Pathology and pathogenesis of the acute MS lesion. *Ms* **12**, 121–132 (2006).
23. Hohlfeld, R. Multiple sclerosis: human model for EAE? *Eur. J. Immunol.* **39**, 2036–2039 (2009).
24. Kaser, A., Zeissig, S. & Blumberg, R. S. Inflammatory bowel disease. *Annu. Rev. Immunol.* **28**, 573–621 (2010).
25. Vaahtovuori, J. *et al.* Fecal microbiota in early rheumatoid arthritis. *J. Rheumatol.* **35**, 1500–1505 (2008).
26. Knip, M. *et al.* Dietary intervention in infancy and later signs of  $\beta$ -cell autoimmunity. *N. Engl. J. Med.* **363**, 1900–1908 (2010).
27. Lauer, K. Diet and multiple sclerosis. *Neurol. (Tokyo)* **49** (suppl. 2), S55–S61 (1997).
28. Kira, J.-I. *et al.* Changes in the clinical phenotypes of multiple sclerosis during the past 50 years in Japan. *J. Neurol. Sci.* **166**, 53–57 (1999).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank I. Arnold-Ammer, N. Reißer and L. Penner for technical assistance, and M. Pfunder, U. Stauffer, C. Hornung and N. Joswig for maintaining our germ-free colony and technical support. We are much obliged to R. Kemler for his support. This work was funded by SFB 571 (Project B6), the German Competence Network on Multiple Sclerosis (KKNMS), ARSEP (France), and by the Max Planck Society. K.B. is supported by a fellowship from ARSEP. Z.A.R. is supported by a PhD fellowship from the Emirates Foundation. M.B. receives a fellowship from the Hellenic Neurological Society.

**Author Contributions** K.B., H.W. and G.K. designed experiments and wrote the manuscript with input from co-authors. K.B. performed most of the experiments. M.M., M.K. and M.B. performed flow cytometry experiments or assisted in experiments. Z.A.R. performed flow cytometry and immunofluorescence staining for brain infiltrates. C.J. supervised the maintenance of germ-free mouse colony and colonization experiments.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to H.W. ([hwekerle@neuro.mpg.de](mailto:hwekerle@neuro.mpg.de)) or G.K. ([guru@neuro.mpg.de](mailto:guru@neuro.mpg.de)).

## METHODS

**Mice and colonization.** Wild-type SJL/J, RR and RR  $\times$  MOG<sup>-/-</sup> SJL/J and IgH<sup>MOG</sup>  $\times$  actin-GFP and actin-GFP SJL/J mice were bred at the animal facility of the Max Planck Institute of Neurobiology. Germ-free RR mice were obtained by transferring embryos, isolated by sterile hysterectomy on embryonic day (E) 18.5, to sterile breeding conditions and by fostering on germ-free foster mothers. They were bred and maintained in positive-pressure plastic isolators and provided with  $\gamma$ -irradiated commercial rodent diet and autoclaved water at the animal facility of the Max Planck Institute of Immunobiology and Epigenetics. Fecal samples were routinely cultured in standard I-Bouillon and examined for contamination. In addition, mice from the colony were screened bi-annually according to FELASA health monitoring recommendations. For re-colonization experiments, germ-free mice were placed in cages with bedding material pre-conditioned by conventional (SPF) mice. All animal procedures were in accordance with the guidelines of the Committee on Animals of the Max Planck Institute of Neurobiology and with a license from the Regierung von Oberbayern.

**Active induction of EAE.** Mice were immunized subcutaneously with 200  $\mu$ g rMOG emulsified in Freund's adjuvant supplemented with 5 mg ml<sup>-1</sup> *Mycobacterium tuberculosis* (strain H37Ra; Difco). On days 0 and 2 after immunization, 200 ng of pertussis toxin (List Biological Laboratories) were injected intraperitoneally. Clinical scoring of EAE was done as published<sup>3</sup>.

**Antibiotic treatment.** For short-term antibiotic treatment, 8-week-old wild-type SJL/J mice were treated for 7 days with 1 g l<sup>-1</sup> of metronidazole (Sigma), 1 g l<sup>-1</sup> of neomycin (Sigma) and 0.5 g l<sup>-1</sup> of vancomycin (AppliChem) in their drinking water.

**CFSE labelling and adoptive transfer.** Splenocytes from RR or wild-type mice were labelled at 37 °C for 10 min with 5  $\mu$ M CFSE (Invitrogen) in PBS containing 1% fetal bovine serum (FBS). Cells were washed twice in ice-cold PBS and subsequently CD4<sup>+</sup> T cells were isolated using a mouse CD4<sup>+</sup> T-cell isolation kit (R&D Systems). 5  $\times$  10<sup>6</sup> CFSE-labelled CD4<sup>+</sup> T cells were injected intravenously into wild-type SJL/J mice.

**B-cell isolation and adoptive transfer.** B cells were isolated from spleens using a mouse B-cell isolation kit (R&D Systems). B cells were enriched to >90% purity, as confirmed by flow cytometry. 10  $\times$  10<sup>6</sup> purified IgH<sup>MOG</sup>-GFP or wild-type-GFP B cells were intravenously injected into RR, wild-type or RR  $\times$  MOG<sup>-/-</sup> mice.

**Proliferation assay.** For the proliferation assay, 2  $\times$  10<sup>5</sup> splenocytes were cultured in the presence of various concentrations of rMOG or anti-CD3 antibody (BD Pharmingen), as indicated. Proliferative response was measured by the incorporation of [<sup>3</sup>H]-thymidine (1  $\mu$ Ci well<sup>-1</sup>) during the last 16 h of a 72 h culture period. Proliferation assays were performed in triplicates.

**Cell isolation and flow cytometry.** Single-cell suspensions were prepared from spleen, pooled peripheral lymph nodes (axillary plus inguinal), or individual lymph nodes (cervical and inguinal), mesenteric lymph nodes or Peyer's patches by mechanical disruption via forcing through 40- $\mu$ m cell strainers (BD Biosciences). For the isolation of lamina propria lymphocytes, small intestine was collected in ice-cold HBSS buffered with 15 mM HEPES. After careful removal

of Peyer's patches, fatty tissue and fecal contents, the intestine was opened longitudinally and cut into small pieces. The intestinal fragments were washed three times for 15 min with stirring (300 r.p.m.) in HBSS containing 5 mM EDTA, 15 mM HEPES and 10% FBS. Next, intestinal pieces were washed once for 5 min with stirring in RPMI containing 15 mM HEPES and 10% FBS, followed by an incubation step at 37 °C with stirring (500 r.p.m.) in RPMI with 15 mM HEPES, 10% FBS and 100 U ml<sup>-1</sup> Collagenase VII (Sigma). The digested tissue was washed once in RPMI with 15 mM HEPES and 10% FBS, before the lamina propria lymphocytes were subjected to FACS analysis. CNS infiltrating cells were purified by Percoll gradient centrifugation as described<sup>3</sup>. For detection of cell surface markers, cells were stained in FACS buffer (PBS containing 1% BSA and 0.1% NaN<sub>3</sub>) with fluorochrome-labelled monoclonal antibodies: PerCP-conjugated anti-CD4 (RM4-5), PerCP-Cy5.5-conjugated anti-B220 (RA3-6B2), PE- and APC-conjugated anti-CD19 (1D3), APC-conjugated anti-CD8 $\alpha$  (53-6.7), PE-conjugated anti-TCR $\gamma\delta$  (eBioGL3), PE-conjugated anti-CD11b (M1/70), APC-conjugated anti-NKp46 (29A1.4), FITC-conjugated anti-CD11c (HL3), biotin-conjugated anti-Gr1 (RB6-8C5), FITC-conjugated anti-CD45.1 (A20), FITC-conjugated anti-V $\beta$ 4 (KT4), PE-conjugated anti-V $\alpha$ 8.3 (B21.14), FITC-conjugated anti-GL7, PE-conjugated anti-Fas (Jo2) and PE-conjugated streptavidin. For intracellular cytokine staining, cells were activated with 50 ng ml<sup>-1</sup> PMA (Sigma) and 500 ng ml<sup>-1</sup> ionomycin (Sigma) in the presence of 5  $\mu$ g ml<sup>-1</sup> brefeldin A (Sigma) for 4 h at 37 °C. After surface staining, cells were fixed and permeabilized in 4% paraformaldehyde/0.1% saponin in HEPES-buffered HBSS and stained intracellularly using the following antibodies: PE-conjugated anti-IL17 (TC11-18H10), APC-conjugated anti-IFN- $\gamma$  (XMG1.2), APC-conjugated anti-TNF- $\alpha$  (MP6-XT22), PE-conjugated anti-IL-10 (JES5-16E3) and APC-conjugated anti-FoxP3 (FJK-16 s). All antibodies were purchased from BD Pharmingen or eBioscience. Cells were acquired on a FACSCalibur (BD Biosciences) and analysis was performed using FlowJo (TreeStar) software.

**ELISA.** Serum titres of anti-MOG antibodies were quantified as previously described<sup>3</sup>. Cytokine levels in cell culture supernatants were determined with antibody pairs for IFN- $\gamma$  (BD Biosciences) or IL-17 (eBioscience).

**Immunofluorescence.** Organs were fixed in PBS with 4% paraformaldehyde and cryoprotected in PBS plus 30% sucrose before embedding in OCT medium (A. Hartenstein). Cryostat sections (10  $\mu$ m in thickness) of spleen, lymph nodes and brains were fixed in acetone. Sections were blocked with PBS and 5% BSA before being stained in a humidified chamber. The following antibodies were used for staining: biotin-conjugated anti-CD4 (BD Pharmingen), purified anti-B220 (BD Pharmingen), biotin-conjugated PNA (Vector Laboratories), Alexa Fluor 568-conjugated anti-rat IgG (Invitrogen), Alexa Fluor 488-conjugated Streptavidin (Invitrogen), APC-conjugated streptavidin (eBioscience) and DAPI (Invitrogen). Images were obtained with a fluorescence microscope (Axiovert 200M; Carl Zeiss) and processed with MetaMorph 7.7 Software and Adobe Photoshop CS4.

**Statistical analysis.** GraphPad Prism 5 (GraphPad Software) was used for all statistical analysis. *P* values <0.05 were considered to be significant.

# Aspartate 112 is the selectivity filter of the human voltage-gated proton channel

Boris Musset<sup>1\*</sup>, Susan M. E. Smith<sup>2\*</sup>, Sindhu Rajan<sup>3</sup>, Deri Morgan<sup>1</sup>, Vladimir V. Cherny<sup>1</sup> & Thomas E. DeCoursey<sup>1</sup>

The ion selectivity of pumps and channels is central to their ability to perform a multitude of functions. Here we investigate the mechanism of the extraordinary selectivity of the human voltage-gated proton channel<sup>1</sup>, H<sub>V</sub>1 (also known as HVCN1). This selectivity is essential to its ability to regulate reactive oxygen species production by leukocytes<sup>2–4</sup>, histamine secretion by basophils<sup>5</sup>, sperm capacitation<sup>6</sup>, and airway pH<sup>7</sup>. The most selective ion channel known, H<sub>V</sub>1 shows no detectable permeability to other ions<sup>1</sup>. Opposing classes of selectivity mechanisms postulate that (1) a titratable amino acid residue in the permeation pathway imparts proton selectivity<sup>1,8–11</sup>, or (2) water molecules ‘frozen’ in a narrow pore conduct protons while excluding other ions<sup>12</sup>. Here we identify aspartate 112 as a crucial component of the selectivity filter of H<sub>V</sub>1. When a neutral amino acid replaced Asp 112, the mutant channel lost proton specificity and became anion-selective or did not conduct. Only the glutamate mutant remained proton-specific. Mutation of the nearby Asp 185 did not impair proton selectivity, indicating that Asp 112 has a unique role. Although histidine shuttles protons in other proteins, when histidine or lysine replaced Asp 112, the mutant channel was still anion-permeable. Evidently, the proton specificity of H<sub>V</sub>1 requires an acidic group at the selectivity filter.

Voltage-gated proton channels are considered specific (perfectly selective) for protons, because no evidence exists for permeation of anything but H<sup>+</sup>. Specificity, combined with a large deuterium isotope effect<sup>9</sup> and extraordinarily strong temperature dependence of conduction<sup>10</sup> suggests a permeation pathway more complex than a simple water wire, as exists in gramicidin<sup>13</sup>. All proton conduction seems consistent with a hydrogen-bonded chain (HBC) mechanism<sup>14</sup>; a HBC including a titratable group could explain several unique properties of H<sub>V</sub>1 (ref. 1), especially proton selectivity<sup>14</sup>. Yet in a recent study, mutation of each titratable amino acid in all four transmembrane helices of H<sub>V</sub>1 failed to abolish conduction<sup>12</sup>. Thus, the mechanism producing proton selectivity remained unknown.

We noticed that a human gene, *C15orf27* (of unknown function), contains a predicted voltage sensor domain (VSD) that shares 25% sequence identity and 52% similarity (<http://www.ebi.ac.uk/Tools/emboss/align/needle>) with the VSD of H<sub>V</sub>1, and includes three Arg residues in the S4 transmembrane helix that are conserved among all known H<sub>V</sub>1 homologues. Phylogenetic analysis of VSD sequences (Supplementary Fig. 1) reveals that a group comprising H<sub>V</sub>1, *C15orf27* and voltage-sensitive phosphatase (VSP) sequences separated early from the two phylogenetically distinct groups of depolarization activated VSDs described previously (K<sub>V</sub> channels and Na<sub>V</sub>/Ca<sub>V</sub> channels), supporting the modular evolution of VSD-containing proteins<sup>15</sup>. Furthermore, H<sub>V</sub>1 VSDs occupy a discrete lineage, distinct from those of VSP and *C15orf27* orthologues.

When we cloned the *C15orf27* gene and expressed the product in HEK-293 or COS-7 cells, the green fluorescent protein (GFP)-tagged protein localized at the plasma membrane (Supplementary Fig. 2), but we detected no currents beyond those in non-transfected cells. We

reasoned that substitutions based on sequence elements that differ between H<sub>V</sub>1 and *C15orf27* should be structurally tolerated while revealing residues responsible for proton conduction. We therefore mutated residues that are perfectly conserved in 21 H<sub>V</sub>1 family members and differ between *C15orf27* and H<sub>V</sub>1. We replaced five candidate residues in H<sub>V</sub>1 (D112, D185, N214, G215 and S219) (Fig. 1a, b) with the corresponding residue in *C15orf27*. Four mutants exhibited large currents under whole-cell voltage clamp (Fig. 1d). The reversal (zero current) potential,  $V_{\text{rev}}$ , measured at several p*H*<sub>o</sub> and p*H*<sub>i</sub> (external and internal pH, respectively), was close to the Nernst potential for protons,  $E_{\text{H}}$  (Fig. 1c), demonstrating proton selectivity. D112V mutants localized to the plasma membrane (Supplementary Fig. 3), but showed no convincing current (Fig. 1d). Some D112V-transfected HEK-293 or COS-7 cells (and non-transfected cells) had small native proton currents. H140A/H193A double mutants<sup>16,17</sup>, in which the two Zn<sup>2+</sup>-binding His residues are neutralized, resemble wild type, with similar Δ*pH*-dependent gating<sup>12</sup>, and  $V_{\text{rev}}$  near  $E_{\text{H}}$  (Supplementary Fig. 4). We expressed mutants in this Zn<sup>2+</sup>-insensitive background (D112X/A/A) to distinguish their currents from native currents that are abolished by 100 μM Zn<sup>2+</sup> at p*H*<sub>o</sub> 7.0. We tentatively concluded that Asp 112 is crucial to proton conduction.

The absence of detectable currents in D112V led us to make other D112X substitutions. These mutants (Fig. 2a) showed slowly activating outward currents upon depolarization that resembled H<sub>V</sub>1 currents. As reported previously<sup>12</sup>, Asp 112 mutation had little effect on the Δ*pH* dependence of gating. The proton conductance–voltage ( $g_{\text{H}}-V$ ) relationship of all D112X mutants shifted roughly –60 mV when p*H*<sub>o</sub> increased from 5.5 to 7.0 (Supplementary Fig. 5), as in wild-type channels<sup>1,8,18</sup>. Mutation of Asp 112 did influence channel opening and closing kinetics (Supplementary Table 2).

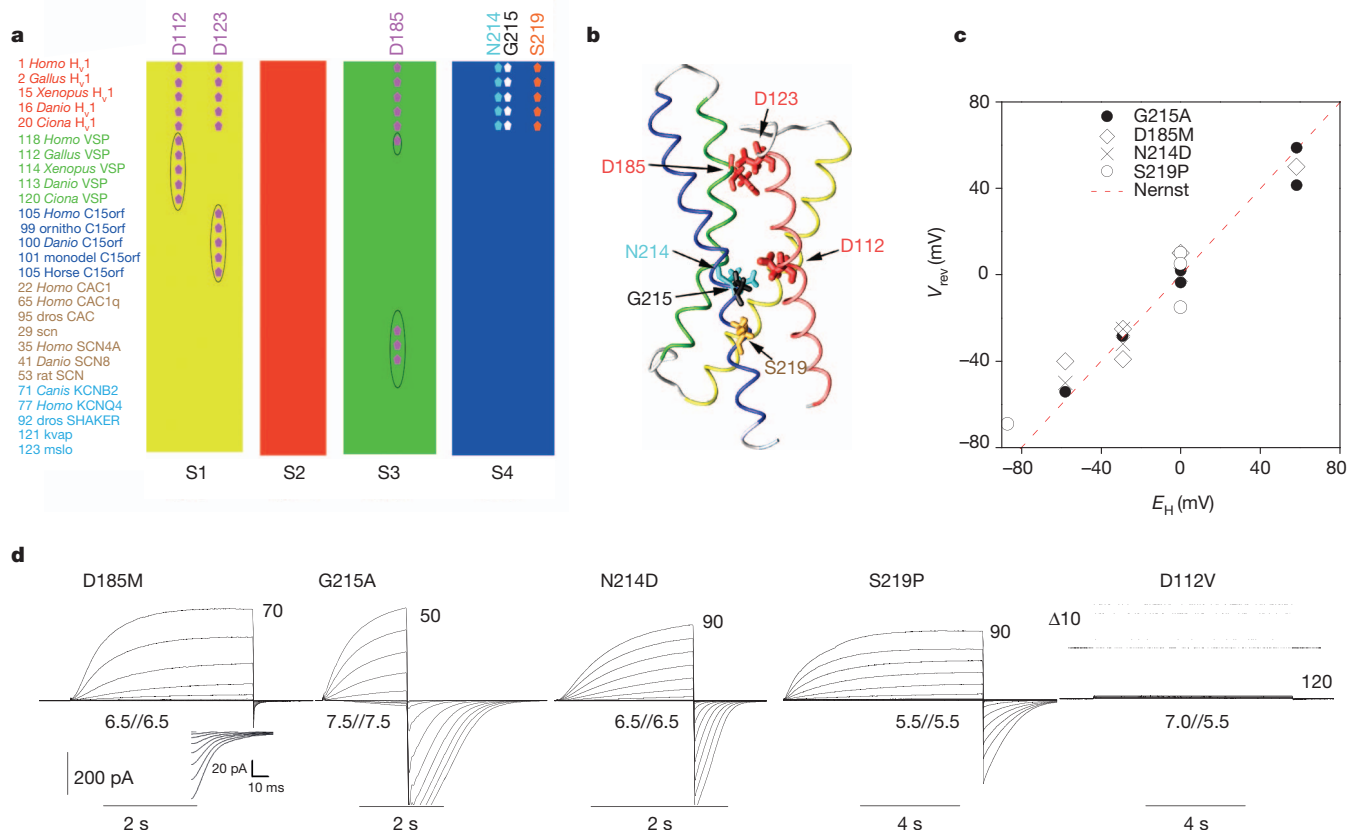
Measurements of  $V_{\text{rev}}$  in Asp 112 mutants showed a marked departure from wild-type H<sub>V</sub>1 properties. At symmetrical p*H* 5.5,  $V_{\text{rev}}$  was near 0 mV (not shown). At p*H*<sub>o</sub> 7.0, p*H*<sub>i</sub> 5.5 (Fig. 2a, column 3), wild-type channels reversed near  $E_{\text{H}}$  (–87 mV), indicating proton selectivity. But for all mutants except D112E,  $V_{\text{rev}}$  was substantially positive to  $E_{\text{H}}$  (Fig. 2b), ranging from –58 mV (D112H) to –13 mV (D112N). Substitutions at Asp 112 eliminated the proton specificity that distinguishes H<sub>V</sub>1 from all other ion channels<sup>1</sup>. A previous study described currents in D112A and D112N mutants<sup>12</sup>, but did not report  $V_{\text{rev}}$ .

We expected that loss of proton selectivity would result in nonselective permeation of cations. Surprisingly,  $V_{\text{rev}}$  did not change detectably when Na<sup>+</sup>, K<sup>+</sup>, *N*-methyl-D-glucamine<sup>+</sup> or TEA<sup>+</sup> (tetraethylammonium<sup>+</sup>) replaced TMA<sup>+</sup> (tetramethylammonium<sup>+</sup>) (Supplementary Table 4). To test anion against cation selectivity, we adopted the classical tactic of replacing a fraction of the bath solution with isotonic sucrose<sup>19</sup>. The Nernst equation predicts that dilution of all extracellular ions except H<sup>+</sup> and OH<sup>–</sup> (leaving internal ion concentrations unchanged) will shift  $V_{\text{rev}}$  negatively for a cation-selective channel, but positively for an anion-selective one. Despite the tenfold reduction of buffer concentration, direct measurement confirmed that p*H* remained constant.

<sup>1</sup>Department of Molecular Biophysics & Physiology, Rush University Medical Center, Chicago, Illinois 60612, USA. <sup>2</sup>Department of Pathology, Emory School of Medicine, Atlanta, Georgia 30322, USA.

<sup>3</sup>Department of Medicine, University of Chicago, Chicago, Illinois 60637, USA.

\*These authors contributed equally to this work.



**Figure 1** | Identification of five key amino acids that differ in  $H_V1$  and C15orf27, and the currents generated in a heterologous expression system by  $H_V1$  mutants in which  $H_V1$  residues were replaced by the corresponding amino acid in the non-conducting C15orf27. **a**, Representative subset of multiple sequence alignment of 122 VSDs, only transmembrane helices are shown. Gene families include  $H_V1$ , voltage-sensitive phosphatases, C15orf27,  $Ca^{2+}$  and  $Na^{+}$  channels, and  $K^{+}$  channels (see Supplementary Fig. 1). **b**, Location of the key amino acids in the open  $H_V1$  channel VSD viewed from the side (membrane), based on a homology model<sup>17</sup>. **c**,  $V_{rev}$  in the four

conducting mutants is near  $E_H$  (dashed line), indicating proton selectivity.  $V_{rev}$  was measured using tail currents; in G215A  $V_{rev}$  was positive to threshold and was observed directly. **d**, Voltage-clamp current families in cells expressing  $H_V1$  mutants. Depolarizing pulses were applied in 10-mV increments from a holding voltage,  $V_{hold} = -40$  mV (D185M, D112V),  $-60$  mV (G215A, S219P), or  $-90$  mV (N214D), with the most positive pulse labelled. After membrane repolarization, an inward 'tail current' is seen as channels close (see inset for D185M); pH is given as  $pH_o/pH_i$ . D112V showed no clear current.

Figure 3 illustrates determination of  $V_{rev}$  from tail currents in a D112H-transfected cell at pH 5.5/5.5 ( $pH_o/pH_i$ ) in methanesulphonate<sup>-</sup> ( $CH_3SO_3^-$ ; Fig. 3a) or  $Cl^-$  solution (Fig. 3c), and after 90% reduction of external ionic strength (Figs 3b, d). Surprisingly, for all Asp 112 mutants, sucrose shifted  $V_{rev}$  positively (Supplementary Fig. 6), indicating anion selectivity both in  $CH_3SO_3^-$  (Fig. 3e) and  $Cl^-$  solutions (Fig. 3f). For  $H_V1$  and D112E,  $V_{rev}$  did not change, reaffirming their proton specificity. Neutralization of a single Aspartate residue converts a proton channel into a predominantly anion selective channel. Thus, Asp 112 mediates charge selectivity as well as proton selectivity.

To confirm anion permeability of Asp 112 mutants, we replaced the main external anion,  $CH_3SO_3^-$ , with  $Cl^-$ . Consistent with previous studies<sup>1</sup>,  $V_{rev}$  in  $H_V1$  was unchanged. As shown in D112H (Fig. 3a, c),  $V_{rev}$  shifted negatively in  $Cl^-$  solutions in all mutants (except D112E), indicating that  $Cl^-$  is more permeant than the larger  $CH_3SO_3^-$  anion (Fig. 3g). That all conducting non-acidic mutants showed  $Cl^-$  permeability indicates that Asp 112 mediates not only proton selectivity, but also charge selectivity. Currents were smaller than wild type in cells expressing some mutants (Supplementary Fig. 7), suggesting a smaller unitary conductance. Evidently, these channels conduct anions, but not very well.

Although the mutant channels have diminished selectivity,  $V_{rev}$  did shift negatively when  $pH_o$  increased from 5.5 to 7.0 (Fig. 2b). Because

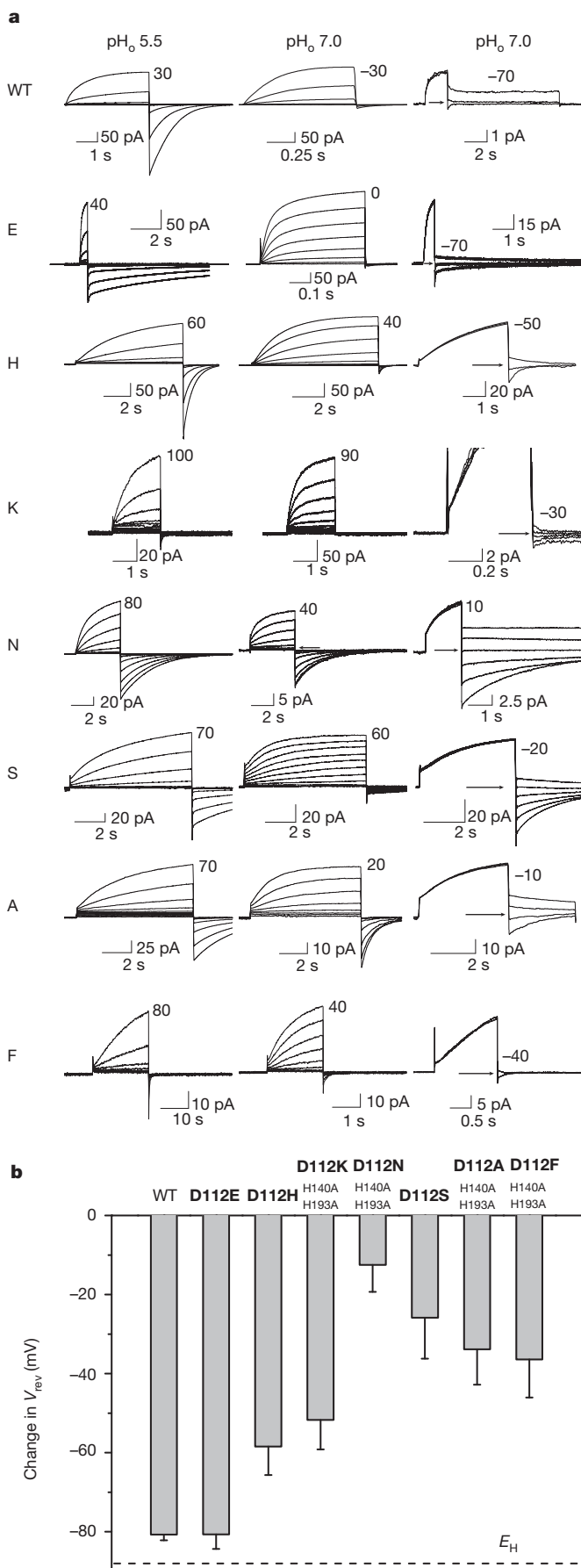
these solutions differ mainly in buffer species and concentrations of  $H^+$  and  $OH^-$ , Asp 112 mutants must have significant permeability to  $H^+$  and/or  $OH^-$ . The Goldman-Hodgkin-Katz equation shows how  $V_{rev}$  depends on ion concentrations:

$$V_{rev} = \frac{RT}{zF} \log \frac{P_{Cl^-} [Cl^-]_i + P_{CH_3SO_3^-} [CH_3SO_3^-]_i + P_{OH^-} [OH^-]_i + P_{H^+} [H^+]_o}{P_{Cl^-} [Cl^-]_o + P_{CH_3SO_3^-} [CH_3SO_3^-]_o + P_{OH^-} [OH^-]_o + P_{H^+} [H^+]_i},$$

where  $R$  is the gas constant,  $T$  the absolute temperature (Kelvin),  $z$  the ionic valence ( $= 1$ ),  $F$  is Faraday's constant and  $P$  is permeability.

Ions with greater permeability dominate  $V_{rev}$ . Permeation of  $H^+$  and  $OH^-$  are difficult to distinguish because they have the same Nernst potential<sup>1</sup>. The data can be interpreted assuming permeation of either (Supplementary Table 3), but the anion selectivity of Asp 112 mutants and the pH dependence of sucrose effects (Figs 3e, f) support  $OH^-$  permeation. The relative permeability of conducting Asp 112 mutants was  $OH^-$  (or  $H^+$ )  $> Cl^- > CH_3SO_3^-$ .

Although Asp 112 is essential to selectivity, other acidic groups might participate. We mutated Asp 185, located in the presumed conduction pore (Fig. 1b)<sup>12,17</sup>. However, like D185M (Fig. 1b), D185V, D185A and D185N remained proton-selective (Supplementary Fig. 8). As evidence against additive effects, the double mutant D112N/D185M did not differ from D112N (Supplementary Fig. 8).



**Figure 2 | Currents in Asp 112 mutants resemble proton currents, but are not.** **a**, Currents generated by wild type (WT), D112E, D112H, D112K/A/A, D112N/A/A, D112S, D112A/A/A and D112F/A/A in COS-7 cells (pH<sub>i</sub> 5.5) at pH<sub>o</sub> 5.5 (column 1) or 7.0 (column 2), during families of pulses in 10 mV increments up to indicated voltages. Tail currents at pH<sub>o</sub> 7.0 (column 3) reveal that  $V_{rev}$  deviates from  $E_{H^+}$ , indicating loss of proton selectivity. At pH<sub>o</sub> 5.5  $V_{hold}$  was  $-40$  mV ( $-60$  mV for WT). At pH<sub>o</sub> 7.0  $V_{hold}$  was  $-40$  mV (K, N, S),  $-50$  mV (F),  $-60$  mV (H, A),  $-80$  mV (E), or  $-90$  mV (WT);  $V_{pre}$  was  $-65$  mV (WT),  $-40$  mV (E),  $-10$  mV (H),  $+50$  mV (K),  $+40$  mV (N, S), or  $+20$  mV (A, F).  $V_{rev}$  (arrows) was determined from the amplitude and direction of tail current decay. For D112N,  $V_{rev}$  was above  $V_{threshold}$  and was evident during pulse families. **b**, Shift in  $V_{rev}$  when the TMACH<sub>3</sub>SO<sub>3</sub> bath solution was changed from pH 5.5 to 7.0. There is no difference between WT and D112E, but the shift in all other mutants is smaller than WT ( $P < 0.001$ , by one-way ANOVA followed by Tukey's test,  $n = 7, 4, 9, 8, 6, 7, 9$  and 4). Error bars in **b** are s.e. Dashed line shows  $E_{H^+}$ .

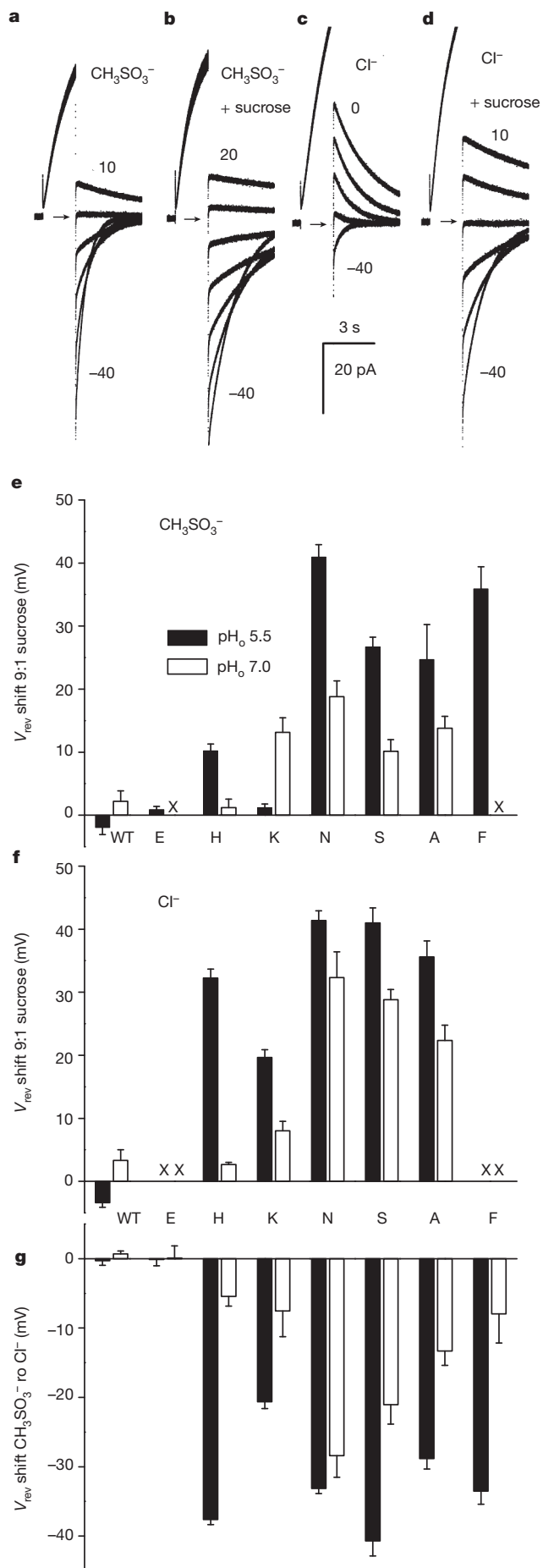
Consistent with earlier predictions that a titratable amino acid provides the selectivity filter of H<sub>v</sub>1 (refs 1, 8–11), only channels with acidic residues (Glu or Asp) at position 112 manifested proton specificity. Asp 112 lies at the constriction of the presumed pore (Fig. 1b), a logical location for a selectivity filter, and just external to the postulated gating charge transfer centre<sup>20</sup>. Our original prediction envisioned selectivity arising from protonation/deprotonation of a residue during conduction, but other mechanisms are possible. For example, proton selectivity of the influenza A M<sub>2</sub> viral proton channel has been explained by (1) immobilized water<sup>21</sup>, (2) successive proton transfer and release by His 37 (refs 22, 23) and (3) delocalization of the proton among His 37 and nearby water molecules<sup>24</sup>.

The  $\text{Cl}^-$  permeability of D112H was completely unexpected, given strong precedents for His imparting proton selectivity to channels. Histidine shuttles protons in  $\text{K}^+$  or  $\text{Na}^+$  channel VSDs with Arg→His mutations<sup>25–27</sup>, in carbonic anhydrase<sup>28</sup> and in  $\text{M}_2$  channels<sup>22,23</sup>. However, these molecules are not proton-specific<sup>27,29</sup>. Evidently, His shuttles protons, but does not guarantee proton selectivity. In  $\text{H}_\text{V}1$ , Asp 112 (or Glu 112 in D112E) excludes anions, resulting in proton-specific conduction. When protonated, Glu and Asp are neutral whereas His is cationic, which may explain why D112H fails to exclude anions.

The anion selectivity of neutral Asp 112 mutants indicates that electrostatic forces due to the charge distribution in the rest of the channel deter cation permeation, and that the cation selectivity of the wild-type channel is due to the anionic charge of Asp 112. Asp 185 does not participate directly in selectivity (Supplementary Fig. 8). VSP family members possess the equivalent of Asp 112 (Fig. 1a), yet conduct no current<sup>30</sup>, illustrating that Asp 112 requires a specific microenvironment to achieve selectivity. Although permeation of  $\text{Cl}^-$  and  $\text{CH}_3\text{SO}_3^-$  suggests a wide pore in D112X mutants, local geometry might differ in wild-type channels due to the presence of anionic Asp 112.

Regulation of voltage gating by  $\Delta\text{pH}$  is distinct from permeation. Pathognomonic of  $\text{H}_{\text{V}1}$  is a strict correlation between the  $g_{\text{H}}-V$  relationship and  $V_{\text{rev}}$ , in which  $V_{\text{threshold}}$  shifts 40 mV per unit change in  $\Delta\text{pH}$  (ref. 8). The  $\Delta\text{pH}$  dependence persisted in mutants with shifted  $g_{\text{H}}-V$  relationships<sup>12</sup>. Here we show uncoupling of  $V_{\text{rev}}$  and voltage gating. Asp 112 mutants retained normal  $\Delta\text{pH}$  dependence (Supplementary Fig. 5), despite the dissociation of  $V_{\text{rev}}$  from  $\Delta\text{pH}$  (Fig. 2 and Supplementary Fig. 9). This uncoupling of pH control of gating from permeation speaks against any mechanism that invokes regulation by local proton concentration in the vicinity of S4 Arg residues<sup>12</sup>.

In summary, Asp 112 is a critical component of the selectivity filter of Hv1, crucial to both proton selectivity and charge selectivity. That D112E was proton-selective, but D112H conducted anions indicates that this proton channel requires an acid at the selectivity filter. That neutralization of nearby Asp 185 did not affect selectivity suggests that Asp 112 has a unique role.



**Figure 3 | Dilution of ionic strength by 90% with isotonic sucrose shifted  $V_{rev}$  positively, indicating that most Asp 112 mutants are anion-selective.** **a**, Measurement of  $V_{rev}$  by tail currents in a cell transfected with D112H at pH 5.5//5.5, **b**, after sucrose. **c**,  $V_{rev}$  in the same cell in pH 5.5  $Cl^-$  solution, and **d**, after sucrose. Arrows indicate zero current.  $V_{hold} = -40$  mV,  $V_{pre} = +60$  mV. **e**, Mean shifts of  $V_{rev}$  with decreasing ionic strength in  $CH_3SO_3^-$  solutions or **f**, in  $Cl^-$  solutions. Each value was determined in 3–6 cells. X = not done. **g**, Shifts of  $V_{rev}$  when  $CH_3SO_3^-$  was replaced by  $Cl^-$ . Values for WT and D112E do not differ significantly from 0 mV. For all anion-selective mutants except D112N, the difference between shifts at pH 5.5 and 7.0 was significant ( $P < 0.001$ , one-way ANOVA followed by Tukey's test;  $n = 3-8$ ). Error bars in **e-g** are s.e.

## METHODS SUMMARY

The pipette solution (also used externally) contained (in mM) 130  $\text{TMACH}_3\text{SO}_3$ , 2  $\text{MgCl}_2$ , 2 EGTA, 80 MES (2-(*N*-morpholino)ethanesulphonic acid), titrated to pH 5.5 with  $\sim 20$  TMAOH. In the pH 5.5 TMAOH solution, TMACl replaced  $\text{TMACH}_3\text{SO}_3$ . Bath solutions at pH 7.0 had (mM) 90  $\text{TMACH}_3\text{SO}_3$  or TMACl, 3  $\text{CaCl}_2$ , 1 EGTA, 100 BES and 36–40 TMAOH. For experiments with  $\text{Zn}^{2+}$ , solutions contained PIPES without EGTA. Experiments were done at 20–25 °C. Currents are shown without leak correction.  $V_{\text{rev}}$  data were corrected for liquid junction potentials measured in each solution<sup>19</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 5 April; accepted 13 September 2011.

Published online 23 October 2011.

1. DeCoursey, T. E. Voltage-gated proton channels and other proton transfer pathways. *Physiol. Rev.* **83**, 475–579 (2003).
2. Capasso, M. *et al.* HVCN1 modulates BCR signal strength via regulation of BCR-dependent generation of reactive oxygen species. *Nature Immunol.* **11**, 265–272 (2010).
3. DeCoursey, T. E., Morgan, D. & Cherny, V. V. The voltage dependence of NADPH oxidase reveals why phagocytes need proton channels. *Nature* **422**, 531–534 (2003).
4. Henderson, L. M., Chappell, J. B. & Jones, O. T. G. The superoxide-generating NADPH oxidase of human neutrophils is electrogenic and associated with an H channel. *Biochem. J.* **246**, 325–329 (1987).
5. Musset, B. *et al.* A pH-stabilizing role of voltage-gated proton channels in IgE-mediated activation of human basophils. *Proc. Natl Acad. Sci. USA* **105**, 11020–11025 (2008).
6. Lishko, P. V., Botchkina, I. L., Fedorenko, A. & Kirichok, Y. Acid extrusion from human spermatozoa is mediated by flagellar voltage-gated proton channel. *Cell* **140**, 327–337 (2010).
7. Iovannisci, D., Illek, B. & Fischer, H. Function of the HVCN1 proton channel in airway epithelia and a naturally occurring mutation, M91T. *J. Gen. Physiol.* **136**, 35–46 (2010).
8. Cherny, V. V., Markin, V. S. & DeCoursey, T. E. The voltage-activated hydrogen ion conductance in rat alveolar epithelial cells is determined by the pH gradient. *J. Gen. Physiol.* **105**, 861–896 (1995).
9. DeCoursey, T. E. & Cherny, V. V. Deuterium isotope effects on permeation and gating of proton channels in rat alveolar epithelium. *J. Gen. Physiol.* **109**, 415–434 (1997).
10. DeCoursey, T. E. & Cherny, V. V. Temperature dependence of voltage-gated H<sup>+</sup> currents in human neutrophils, rat alveolar epithelial cells, and mammalian phagocytes. *J. Gen. Physiol.* **112**, 503–522 (1998).
11. DeCoursey, T. E. & Cherny, V. V. Voltage-activated hydrogen ion currents. *J. Membr. Biol.* **141**, 203–223 (1994).
12. Ramsey, I. S. *et al.* An aqueous H<sup>+</sup> permeation pathway in the voltage-gated proton channel Hv1. *Nature Struct. Mol. Biol.* **17**, 869–875 (2010).
13. Levitt, D. G., Elias, S. R. & Hautman, J. M. Number of water molecules coupled to the transport of sodium, potassium and hydrogen ions via gramicidin, nonactin or valinomycin. *Biochim. Biophys. Acta* **512**, 436–451 (1978).
14. Nagle, J. F. & Morowitz, H. J. Molecular mechanisms for proton transport in membranes. *Proc. Natl Acad. Sci. USA* **75**, 298–302 (1978).
15. Nelson, R. D., Kuan, G., Saier, M. H. Jr & Montal, M. Modular assembly of voltage-gated channel proteins: a sequence analysis and phylogenetic study. *J. Mol. Microbiol. Biotechnol.* **1**, 281–287 (1999).
16. Ramsey, I. S., Moran, M. M., Chong, J. A. & Clapham, D. E. A voltage-gated proton-selective channel lacking the pore domain. *Nature* **440**, 1213–1216 (2006).
17. Musset, B. *et al.* Zinc inhibition of monomeric and dimeric proton channels suggests cooperative gating. *J. Physiol. (Lond.)* **588**, 1435–1449 (2010).
18. Musset, B. *et al.* Detailed comparison of expressed and native voltage-gated proton channel currents. *J. Physiol. (Lond.)* **586**, 2477–2486 (2008).
19. Barry, P. H. The reliability of relative anion-cation permeabilities deduced from reversal (dilution) potential measurements in ion channel studies. *Cell Biochem. Biophys.* **46**, 143–154 (2006).
20. Tao, X., Lee, A., Limapichat, W., Dougherty, D. A. & MacKinnon, R. A gating charge transfer center in voltage sensors. *Science* **328**, 67–73 (2010).

21. Sansom, M. S. P., Kerr, I. D., Smith, G. R. & Son, H. S. The influenza A virus M2 channel: a molecular modeling and simulation study. *Virology* **233**, 163–173 (1997).
22. Hu, F., Luo, W. & Hong, M. Mechanisms of proton conduction and gating in influenza M2 proton channels from solid-state NMR. *Science* **330**, 505–508 (2010).
23. Venkataraman, P., Lamb, R. A. & Pinto, L. H. Chemical rescue of histidine selectivity filter mutants of the M2 ion channel of influenza A virus. *J. Biol. Chem.* **280**, 21463–21472 (2005).
24. Acharya, R. *et al.* Structure and mechanism of proton transport through the transmembrane tetrameric M2 protein bundle of the influenza A virus. *Proc. Natl Acad. Sci. USA* **107**, 15075–15080 (2010).
25. Starace, D. M. & Bezanilla, F. Histidine scanning mutagenesis of basic residues of the S4 segment of the *Shaker* K<sup>+</sup> channel. *J. Gen. Physiol.* **117**, 469–490 (2001).
26. Starace, D. M. & Bezanilla, F. A proton pore in a potassium channel voltage sensor reveals a focused electric field. *Nature* **427**, 548–553 (2004).
27. Sokolov, S., Scheuer, T. & Catterall, W. A. Ion permeation and block of the gating pore in the voltage sensor of Nav1.4 channels with hypokalemic periodic paralysis mutations. *J. Gen. Physiol.* **136**, 225–236 (2010).
28. Tu, C. K., Silverman, D. N., Forsman, C., Jonsson, B. H. & Lindskog, S. Role of histidine 64 in the catalytic mechanism of human carbonic anhydrase II studied with a site-specific mutant. *Biochemistry* **28**, 7913–7918 (1989).
29. Leiding, T., Wang, J., Martinsson, J., DeGrado, W. F. & Årsköld, S. P. Proton and cation transport activity of the M2 proton channel from influenza A virus. *Proc. Natl Acad. Sci. USA* **107**, 15409–15414 (2010).
30. Murata, Y., Iwasaki, H., Sasaki, M., Inaba, K. & Okamura, Y. Phosphoinositide phosphatase activity coupled to an intrinsic voltage sensor. *Nature* **435**, 1239–1243 (2005).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank P. H. Barry, D. Gillespie, V. S. Markin, J. F. Nagle, R. Pomès, D. Silverman and V. Sokolov for discussions or comments on the manuscript. Supported by NSF grant MCB-0943362 (S.M.E.S. and T.E.D.) and NIH grant GM087507 (T.E.D.). The content is solely the responsibility of the authors and does not necessarily represent the views of the National Institute of General Medical Sciences or the National Institutes of Health.

**Author Contributions** S.R. identified the similarity of C15orf27 to H<sub>v</sub>1 and cloned the *C15orf27* gene; S.M.E.S. conceived the strategic approach based on molecular model, sequence and phylogenetic analysis; S.R. and S.M.E.S. created mutants; T.E.D., B.M. and V.V.C. designed experiments; B.M., D.M. and V.V.C. recorded, analysed and interpreted data; T.E.D. wrote the manuscript; all authors read and approved the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to T.E.D. ([tdecours@rush.edu](mailto:tdecours@rush.edu)).

## METHODS

Exhaustive searches to identify H<sub>V</sub>1 homologues were performed using protein BLAST and PSI-BLAST. A sample of VSDs from K<sup>+</sup>, Na<sup>+</sup> and Ca<sup>2+</sup> channels (that open with depolarization like H<sub>V</sub>1, and in addition one that opens with hyperpolarization), along with putative H<sub>V</sub>1, VSP and C15orf27 homologues were chosen. For cation channels, we sampled from the range of subfamilies, from the VSD repeats within Na<sup>+</sup> and Ca<sup>2+</sup> channels, and from the range of species. VSD sequences, including crystallized K<sup>+</sup> channels (PDB accessions 1ORS, 2R9R and 2A79), were aligned using Promals3D<sup>31</sup>, which incorporates structural information, allowing high-confidence identification of VSD boundaries. Sequences were trimmed to the VSD, realigned with Promals3D, and the resulting alignment was analysed with PhyML (maximum likelihood)<sup>32</sup> and Protpars (maximum parsimony)<sup>33</sup> at the Mobyle portal<sup>34</sup>. Trees were visualized with TreeDyn<sup>35</sup> and iTOL<sup>36</sup>. Parsimony (not shown) and maximum likelihood trees had similar topology, including H<sub>V</sub>1 and C15orf27 families separating into discrete branches. A homology model of the VSD of H<sub>V</sub>1 was constructed as described previously<sup>17</sup>.

The C15orf27 clone was PCR-amplified from human cerebellum and subcloned into pcDNA3.1(+) expression vector (Invitrogen). The coding sequence of human H<sub>V</sub>1 (*HVCN1*) was cloned into either pcDNA3.1(−) or pQBI25-fC3 (to make GFP–H<sub>V</sub>1) vectors as described previously<sup>16</sup>. Site-directed mutants were created using the Stratagene QuikChange (Agilent) procedure according to the manufacturer's instructions. All the positive clones were sequenced to confirm the presence of the introduced mutation. HEK-293 or, more often COS-7 cells were grown to ~80% confluency in 35-mm cultures dishes, usually by seeding cells 1 day ahead of transfection. Cells were transfected with 0.4–0.5 µg of the appropriate cDNA using Lipofectamine 2000 (Invitrogen). After 6 h at 37 °C in 5% CO<sub>2</sub>,

the cells were trypsinized and replated onto glass cover slips at low density for patch clamp recording the following day. We selected green cells under fluorescence for recording. Patch-clamp methods were described previously<sup>18</sup>.

The main pipette solution (also used externally) contained (in mM) 130 TMACH<sub>3</sub>SO<sub>3</sub>, 2 MgCl<sub>2</sub>, 2 EGTA, 80 MES, titrated to pH 5.5 with ~20 TMAOH. In the pH 5.5 TMACl solution, TMACl replaced TMACH<sub>3</sub>SO<sub>3</sub>. Bath solutions at pH 7.0 had (mM) 90 TMACH<sub>3</sub>SO<sub>3</sub> or TMACl, 3 CaCl<sub>2</sub>, 1 EGTA, 100 BES, and 36–40 TMAOH. For experiments with Zn<sup>2+</sup>, solutions contained PIPES buffer<sup>37</sup> without EGTA. Experiments were done at 20–25 °C. Currents are shown without leak correction. *V*<sub>rev</sub> data were corrected for liquid junction potentials measured in each solution<sup>19</sup>.

31. Pei, J., Kim, B. H. & Grishin, N. V. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300 (2008).
32. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
33. Felsenstein, J. PROTPARS – Protein Sequence Parsimony Method <http://cmgm.stanford.edu/phyliip/protpars.html> (1993).
34. Néron, B. *et al.* Mobyle: a new full web bioinformatics framework. *Bioinformatics* **25**, 3005–3011 (2009).
35. Chevenet, F., Brun, C., Bañuls, A. L., Jacq, B. & Christen, R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* **7**, 439 (2006).
36. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).
37. Cherny, V. V. & DeCoursey, T. E. pH-dependent inhibition of voltage-gated H<sup>+</sup> currents in rat alveolar epithelial cells by Zn<sup>2+</sup> and other divalent cations. *J. Gen. Physiol.* **114**, 819–838 (1999).

# Lowland–upland migration of sauropod dinosaurs during the Late Jurassic epoch

Henry C. Fricke<sup>1</sup>, Justin Henceroth<sup>1</sup> & Marie E. Hoerner<sup>1†</sup>

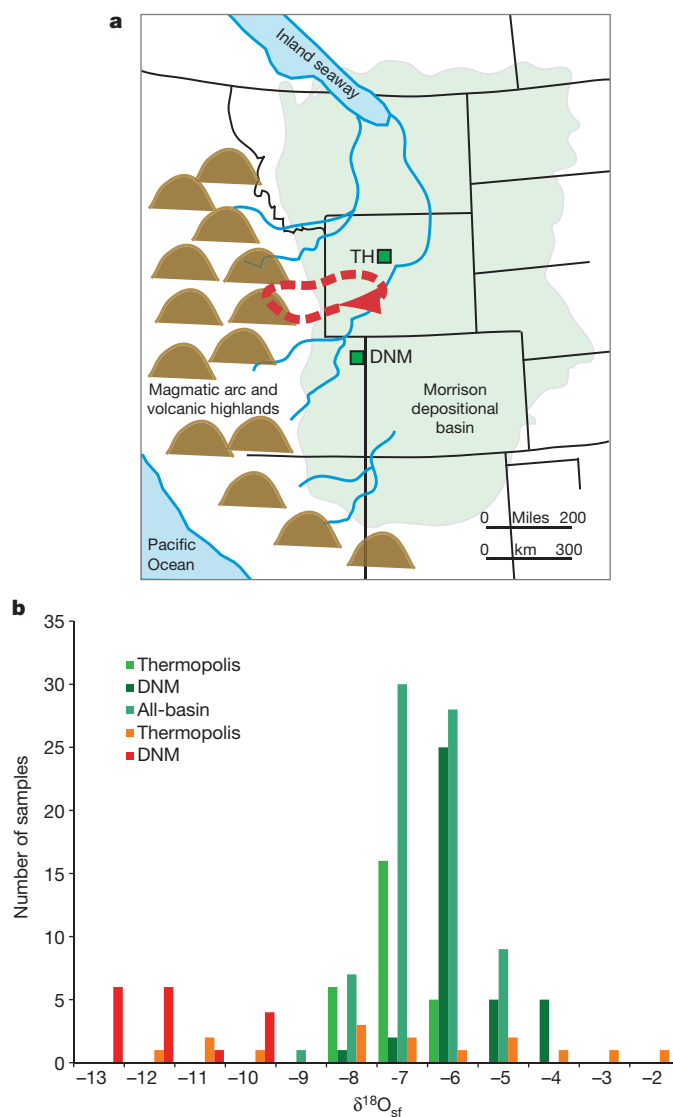
Sauropod dinosaurs were the largest vertebrates ever to walk the Earth, and as mega-herbivores they were important parts of terrestrial ecosystems. In the Late Jurassic-aged Morrison depositional basin of western North America, these animals occupied lowland river-floodplain settings characterized by a seasonally dry climate<sup>1,2</sup>. Massive herbivores with high nutritional and water needs could periodically experience nutritional and water stress under these conditions, and thus the common occurrence of sauropods in this basin has remained a paradox. Energetic arguments and mammalian analogues have been used to suggest that migration allowed sauropods access to food and water resources over a wide region or during times of drought or both<sup>3,4</sup>, but there has been no direct support for these hypotheses. Here we compare oxygen isotope ratios ( $\delta^{18}\text{O}$ ) of tooth-enamel carbonate from the sauropod *Camarasaurus* with those of ancient soil, lake and wetland (that is, 'authigenic') carbonates that formed in lowland settings. We demonstrate that certain populations of these animals did in fact undertake seasonal migrations of several hundred kilometres from lowland to upland environments. This ability to describe patterns of sauropod movement will help to elucidate the role that migration played in the ecology and evolution of gigantism of these and associated dinosaurs.

Inferring the behaviour of ancient organisms is difficult, but geochemical information preserved in their fossil remains can provide such an opportunity. This study of sauropod dinosaur behaviour relies on the fact that  $\delta^{18}\text{O}$  values of surface waters ( $\delta^{18}\text{O}_{\text{sf}}$  for example streams, lakes) vary significantly over any given landscape in response to differences in aridity and elevation among other environmental factors<sup>5,6</sup>. Authigenic carbonates ( $\text{CaCO}_3$ ) form in basin soils, lakes and wetlands, and record the oxygen isotopic characteristics of these host isotopic domains when they precipitate. Similarly vertebrate tooth enamel (bioapatite  $\text{Ca}_5(\text{PO}_4)_3(\text{OH}, \text{CO}_3)$ ) records the oxygen isotope characteristics of the surface water reservoirs that serve as their drinking water<sup>7,8</sup>. If  $\delta^{18}\text{O}_{\text{sf}}$  inferred from 'non-migratory' authigenic carbonates and from dinosaur tooth enamel differ, then it can be concluded that dinosaurs were drinking water that fell outside the basin and thus they travelled outside it.

To use this approach we analysed enamel carbonate from teeth ( $n = 32$ ) of *Camarasaurus* sp. and *Camarasaurus lentus* collected at Thermopolis, Wyoming, and Dinosaur National Monument, Utah (DNM), respectively (Fig. 1a). Palaeosol and lacustrine carbonates were also analysed from DNM ( $n = 38$ ; see Supplementary Information for details on methods and statistics). In addition, we used published  $\delta^{18}\text{O}$  data obtained from a variety of authigenic carbonates found over the entire Morrison basin including the Thermopolis area<sup>9–12</sup>. Comparisons of isotopic data from co-occurring authigenic carbonates and tooth enamel, from tooth-enamel carbonate and tooth-enamel phosphate, and from single teeth indicate that primary palaeobiological information is preserved in tooth enamel (see Supplementary Information for more details about diagenesis).

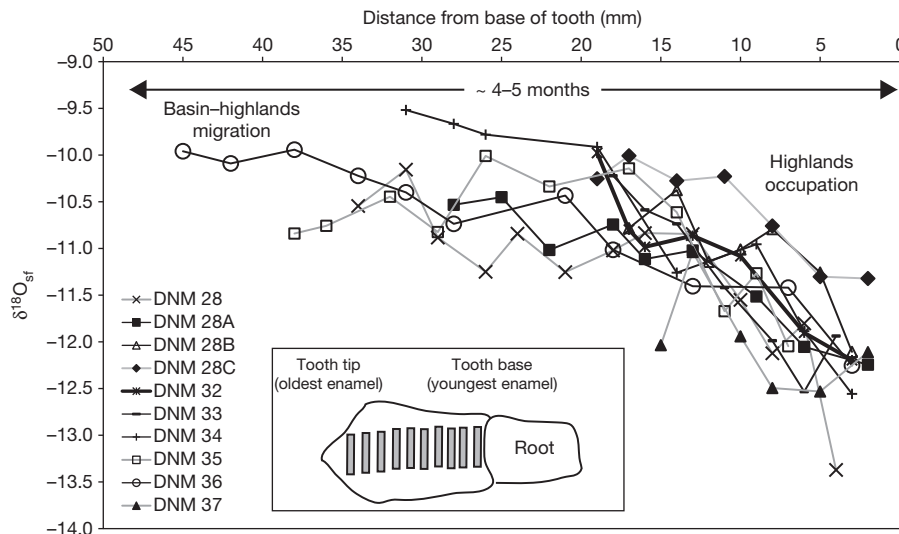
To estimate  $\delta^{18}\text{O}_{\text{sf}}$  using dinosaur tooth enamel, it is assumed that they fractionated oxygen isotopes in a manner similar to all water-dependent

vertebrates studied so far, including birds, mammals and reptiles<sup>7,8</sup>. To estimate  $\delta^{18}\text{O}_{\text{sf}}$  using authigenic carbonate, it is assumed that oxygen isotope fractionation occurred at 24 °C, a temperature consistent with modelled mean annual temperature for the region<sup>13</sup> (see



**Figure 1** | Fossil localities, inferred oxygen isotope ratios of surface water and possible *Camarasaurus* migration routes. **a**, Palaeogeography of western North America during late Jurassic/Morrison time (after refs 2, 9), including fossil localities and one hypothetical migration route. **b**,  $\delta^{18}\text{O}_{\text{sf}}$  estimated using tooth enamel (reds) and authigenic carbonates (greens; Thermopolis data from ref. 11; all-basin data from refs 9, 10, 12). See text and Supplementary Information for details.

<sup>1</sup>Department of Geology, Colorado College, Colorado Springs, Colorado 80903, USA. <sup>†</sup>Present address: Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA.



**Figure 2 | Oxygen isotope ratios of serial enamel samples compared with position relative to the base of the tooth for ten different teeth from DNM.** Teeth form incrementally such that the oldest enamel is at the tip and the youngest enamel is at the base (see inset). Most teeth are characterized by a

gradual decrease in oxygen isotope ratios over time. Tooth DNM 36 is unworn and thus preserves the longest temporal record (approximately 4–5 months), whereas records from other worn teeth are truncated to various degrees.

Supplementary Information for details about estimates of  $\delta^{18}\text{O}_{\text{sf}}$ . Although  $\delta^{18}\text{O}_{\text{sf}}$  inferred for all authigenic carbonates have relatively high and limited  $\delta^{18}\text{O}$  values, those inferred for *Camarasaurus* have significantly different means and variances (Supplementary Tables 1–3). Most importantly, they preserve a record of the lowest  $\delta^{18}\text{O}_{\text{sf}}$  of all carbonates analysed from the Morrison basin (Fig. 1b).

$\delta^{18}\text{O}_{\text{sf}}$  values inferred from *Camarasaurus* that overlap with those from authigenic carbonates are consistent with these animals spending time in the fluvial and wetland environments of the basin. In contrast, lower  $\delta^{18}\text{O}$  of surface water and precipitation ( $\delta^{18}\text{O}_{\text{sf}} < \text{approximately } -9\text{‰}$ ) implied by a large proportion of *Camarasaurus* teeth indicates that they occupied non-basinal settings. Low  $\delta^{18}\text{O}_{\text{sf}}$  values result from the preferential rainout of  $^{18}\text{O}$  from air masses as they rise, cool and lose water while crossing topographic barriers such as volcanic highlands west of the basin (Fig. 1a). Thus, they indicate that animals from both DNM and Thermopolis were drinking water from these high-elevation regions. Although it is possible that ‘extra-basinal’ high-elevation waters could flow into lakes and rivers located in the basin proper, the fact that lake and wetland carbonates do not have low  $\delta^{18}\text{O}$  indicates that such recharge did not have a major influence on  $\delta^{18}\text{O}_{\text{sf}}$  in the basin (Fig. 1b). Therefore *Camarasaurus* populations in these areas must have directly occupied high-elevation regions for at least part of the year before returning to the basin where they died. To do so, these animals must have migrated approximately 300 km in each direction based on palaeogeographical reconstructions for the Late Jurassic<sup>2,9</sup> (Fig. 1a).

Patterns in  $\delta^{18}\text{O}$  obtained from single teeth provide evidence that this migration was seasonal in nature. Because vertebrate teeth, including those of dinosaurs<sup>14</sup>, form incrementally, sequential sampling along the length of a tooth provides a record of  $\delta^{18}\text{O}_{\text{sf}}$  ingested during the time of tooth formation<sup>15–17</sup>. Intra-tooth variations in  $\delta^{18}\text{O}_{\text{sf}}$  inferred from camarasaurid teeth of a single DNM *C. lentus* skull (see Supplementary Information) appear to capture slightly less than half of the sinusoidal cycle that is expected for a single year<sup>15,16</sup>, thus indicating that these teeth formed over approximately 4–5 months (Fig. 2). The specific pattern in  $\delta^{18}\text{O}_{\text{sf}}$  implies that this animal moved out of the basin into highland regions over the period of tooth formation, yet the teeth are found in the basin. Such a situation is possible because tooth enamel does not provide an instantaneous record of ingested  $\delta^{18}\text{O}_{\text{sf}}$  rather, there is a temporal lag associated with the turnover of oxygen in the body. This lag is of the order of 2 weeks for small mammals<sup>18</sup>, and although the length of time is unknown for

sauropod dinosaurs it cannot have been longer than several weeks to a month, otherwise seasonal variations would be obscured altogether. Thus, over 5–6 months, this individual left the basin for the highlands and then returned to the DNM area.

Assuming that *Camarasaurus* migrated in an effort to obtain the food and water they needed to survive, they would have left the basin during the dry season (presumably summer<sup>19</sup>) when plant growth was limited and drought might have been common, and then returned in the wet season (presumably winter<sup>19</sup>). The fact that the DNM *C. lentus* died before preserving a record of basin  $\delta^{18}\text{O}_{\text{sf}}$  in its tooth enamel suggests that it was recently returned, and that it died during the transition from the dry to wet season. The similarity in bulk  $\delta^{18}\text{O}$  among DNM teeth from other individuals suggests that other *Camarasaurus* from DNM exhibited similar behaviour. Without well-constrained intra-tooth data from Thermopolis it is not possible to describe *Camarasaurus* migrational patterns in as much detail in this area. However, the fact that  $\delta^{18}\text{O}_{\text{sf}}$  inferred from Thermopolis are generally higher than those from DNM (Fig. 1b) could mean that the Thermopolis teeth captured a different part of the seasonal cycle in  $\delta^{18}\text{O}_{\text{sf}}$  and thus might have died during a different time of the year, that Thermopolis teeth grew during different year(s), or that these animals visited a different (possibly lower elevation) part of the western highlands.

Overall, the research presented here provides strong support for the hypothesis that *Camarasaurus* could undertake long seasonal migrations. It does not, however, imply that they must have done so. Ongoing studies of other *Camarasaurus* populations and of other sauropods living in different areas will allow us to determine if migrations were a universal characteristic of these animals, or whether it was a behavioural response to environmental stress. In turn it will be possible to address the role that migration might have played in the evolution of sauropod gigantism.

Received 8 August; accepted 19 September 2011.

Published online 26 October 2011.

1. Parrish, J., Peterson, F. & Turner, C. E. Jurassic ‘savannah’ – plant taphonomy and climate of the Morrison Formation (Upper Jurassic, Western USA). *Sedim. Geol.* **167**, 139–164 (2004).
2. Turner, C. E. & Peterson, F. Reconstruction of the Upper Jurassic Morrison Formation extinct ecosystem—a synthesis. *Sedim. Geol.* **167**, 309–355 (2004).
3. Englemann, G. F., Chure, D. J. & Fiorillo, A. R. The implications of a dry climate for the paleoecology of the fauna of the Upper Jurassic Morrison Formation. *Sedim. Geol.* **167**, 297–308 (2004).

4. Dodson, P., Behrensmeyer, A. K., Bakker, R. T. & McIntosh, J. S. Taphonomy and paleoecology of the dinosaur beds of the Jurassic Morrison Formation. *Paleobiology* **6**, 208–232 (1980).
5. Gat, J. R. Oxygen and hydrogen isotopes in the hydrologic cycle. *Annu. Rev. Earth Planet. Sci.* **24**, 225–262 (1996).
6. Bowen, G. Isoscapes: spatial pattern in isotopic biogeochemistry. *Annu. Rev. Earth Planet. Sci.* **38**, 161–187 (2010).
7. Kohn, M. J. & Cerling, T. E. Stable isotope compositions of biological apatite. *Rev. Mineral. Geochem.* **48**, 455–488 (2002).
8. Amiot, R. *et al.* Latitudinal temperature gradient during the Cretaceous Upper Campanian–Middle Maastrichtian:  $\delta^{18}\text{O}$  record of continental vertebrates. *Earth Planet. Sci. Lett.* **226**, 255–272 (2004).
9. Dunegan, S. & Turner, C. Regional paleohydrologic and paleoclimatic settings of wetland/lacustrine depositional systems in the Morrison Formation (Upper Jurassic), Western Interior, USA. *Sedim. Geol.* **167**, 269–298 (2004).
10. Lovelace, D. An Upper Jurassic Morrison Formation fire-induced debris flow: taphonomy and paleoenvironment of a sauropod (*Sauropoda: Supersaurus vivianae*) locality, east-central Wyoming. *Bull. New Mexico Mus. Nat. Hist.* **36**, 47–56 (2006).
11. Jennings, D. S. & Lovelace, D. M. Driese, S.G. Differentiating paleowetland subenvironments using a multi-disciplinary approach: an example from the Morrison Formation, South Central Wyoming, USA. *Sedim. Geol.* 10.1016/j.sedgeo.2011.03.005 (2011).
12. Fallin, M. J. *Carbon Isotope Stratigraphy of the Morrison and Cloverly Formations and Assessment of Vertical Color Change in the Morrison Formation, Coyote Basin, Wyoming*. Master's thesis, Univ. Colorado, Boulder (2005).
13. Sellwood, B. W. & Valdes, P. J. Jurassic climates. *Proc. Geol. Assoc.* **119**, 5–17 (2008).
14. Erickson, G. M. Incremental lines of von Ebner and the assessment of tooth replacement rates using growth line counts. *Proc. Natl Acad. Sci. USA* **93**, 14623–14627 (1996).
15. Fricke, H. C. & O'Neil, J. R. Inter- and intra-tooth variation in the oxygen isotope composition of mammalian tooth enamel phosphate: implications for palaeoclimatological and palaeobiological research. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **126**, 91–99 (1996).
16. Kohn, M. J., Schoeninger, M. J. & Valley, J. W. Variability in oxygen compositions of herbivore teeth: reflections of seasonality or developmental physiology? *Chem. Geol.* **152**, 97–112 (1998).
17. Stanton-Thomas, K. J. & Carlson, S. J. Microscale  $\delta^{18}\text{O}$  and  $\delta^{13}\text{C}$  isotopic analysis of an ontogenetic series of the hadrosaurid dinosaur *Edmontosaurus*: implications for physiology and ecology. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **206**, 257–287 (2004).
18. Podlesak, D. W. *et al.* Turnover of oxygen and hydrogen isotopes in the body water,  $\text{CO}_2$ , hair, and enamel of a small mammal. *Geochim. Cosmochim. Acta* **72**, 19–35 (2008).
19. Rees, P. M., Ziegler, A. M. & Valdes, P. J. in *Warm Climates in Earth History* (eds Huber, B. T. *et al.*) 297–318 (Cambridge Univ. Press, 2003).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank S. Hartman, D. Lovelace and D. Chure for providing context and fossil specimens. The research was funded in part by National Science Foundation EAR-0319041 and Colorado College. We thank G. Olack and A. Colman for the phosphate oxygen isotope analyses, supported by National Science Foundation Major Research Instrumentation award EAR 0923831.

**Author Contributions** J.H. analysed samples and interpreted data from Dinosaur National Monument; M.E.H. analysed samples, interpreted data from Thermopolis and did the statistical analysis. H.C.F. designed the study, aided in data interpretation and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to H.C.F. ([hfricke@coloradocollege.edu](mailto:hfricke@coloradocollege.edu)).

# Experimental infection of bats with *Geomyces destructans* causes white-nose syndrome

Jeffrey M. Lorch<sup>1,2</sup>, Carol U. Meteyer<sup>2</sup>, Melissa J. Behr<sup>3</sup>, Justin G. Boyles<sup>4</sup>, Paul M. Cryan<sup>5</sup>, Alan C. Hicks<sup>6</sup>, Anne E. Ballmann<sup>2</sup>, Jeremy T. H. Coleman<sup>7</sup>, David N. Redell<sup>8</sup>, DeeAnn M. Reeder<sup>9</sup> & David S. Blehert<sup>2</sup>

White-nose syndrome (WNS) has caused recent catastrophic declines among multiple species of bats in eastern North America<sup>1,2</sup>. The disease's name derives from a visually apparent white growth of the newly discovered fungus *Geomyces destructans* on the skin (including the muzzle) of hibernating bats<sup>1,3</sup>. Colonization of skin by this fungus is associated with characteristic cutaneous lesions that are the only consistent pathological finding related to WNS<sup>4</sup>. However, the role of *G. destructans* in WNS remains controversial because evidence to implicate the fungus as the primary cause of this disease is lacking. The debate is fuelled, in part, by the assumption that fungal infections in mammals are most commonly associated with immune system dysfunction<sup>5–7</sup>. Additionally, the recent discovery that *G. destructans* commonly colonizes the skin of bats of Europe, where no unusual bat mortality events have been reported<sup>8–10</sup>, has generated further speculation that the fungus is an opportunistic pathogen and that other unidentified factors are the primary cause of WNS<sup>11,12</sup>. Here we demonstrate that exposure of healthy little brown bats (*Myotis lucifugus*) to pure cultures of *G. destructans* causes WNS. Live *G. destructans* was subsequently cultured from diseased bats, successfully fulfilling established criteria for the determination of *G. destructans* as a primary pathogen<sup>13</sup>. We also confirmed that WNS can be transmitted from infected bats to healthy bats through direct contact. Our results provide the first direct evidence that *G. destructans* is the causal agent of WNS and that the recent emergence of WNS in North America may represent translocation of the fungus to a region with a naive population of animals<sup>8</sup>. Demonstration of causality is an instrumental step in elucidating the pathogenesis<sup>14</sup> and epidemiology<sup>15</sup> of WNS and in guiding management actions to preserve bat populations against the novel threat posed by this devastating infectious disease.

To test the ability of *G. destructans* to act as a primary pathogen, we housed healthy little brown bats (*Myotis lucifugus*;  $n = 29$ ) in the laboratory under hibernation conditions and treated them with conidia of *G. destructans* harvested from pure culture. Histological examination of treated bats that died during the course of the experiment showed that lesions diagnostic for WNS were apparent by 83 days after treatment. All treated bats were positive for WNS by histology when the trial was terminated at 102 days after treatment. In contrast, at the end of the experiment, all bats from the negative control group (bats treated identically but not exposed to conidia of *G. destructans*;  $n = 34$ ) were negative for WNS by histology.

We also investigated the potential for WNS to be transmitted from infected to healthy animals by co-housing hibernating bats naturally infected with WNS (collected from an affected hibernaculum and showing clinical signs of the disease;  $n = 25$ ) with healthy bats (contact exposure group;  $n = 18$ ). Eighty-nine per cent of bats in the contact exposure group developed WNS lesions by day 102, demonstrating for the first time that WNS is transmissible. This has important

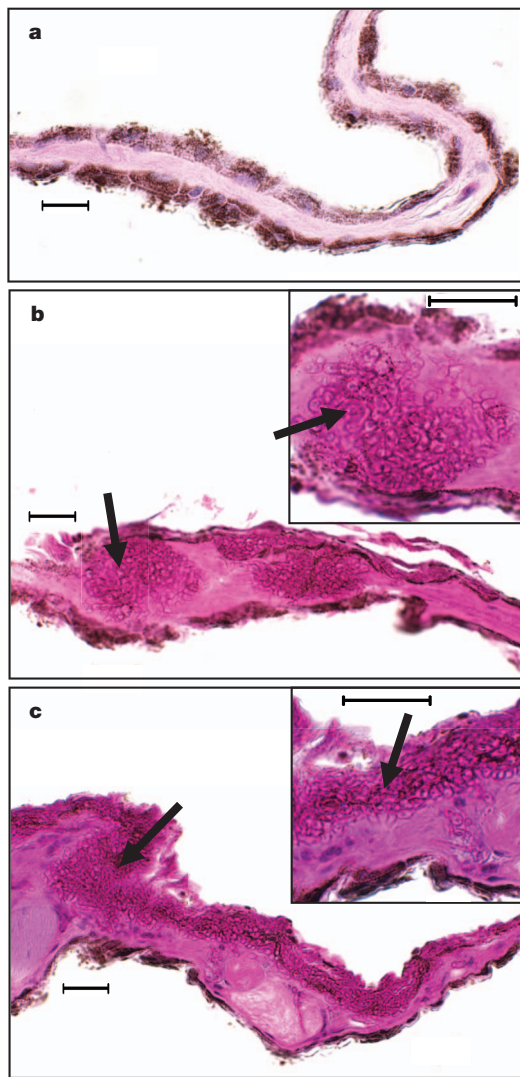
epidemiological and disease management implications, because many of the bat species most commonly impacted by WNS often form tight, occasionally mixed-species clusters during hibernation, facilitating the transfer of fungus among individuals and species. In addition, bat species affected by WNS engage in 'swarming' behaviour at hibernacula just before hibernation. During this time, there is much direct contact between individuals as they participate in a promiscuous mating system<sup>16</sup>. Furthermore, individual bats have been documented to move long distances between hibernacula during this period<sup>17</sup>, which may, in part, facilitate the spread of WNS across the landscape.

To determine if WNS could be spread between bats through the air, healthy bats ( $n = 36$ ) were placed in mesh cages in close proximity to (separated by 1.3 cm), but not in direct contact with, the positive control and treated groups. After a period of 102 days, none of the animals exposed to possible airborne conidia from bats with WNS showed histopathological evidence of infection. This may be due to an inability of *G. destructans* conidia to travel through air at levels sufficient to establish infections in neighbouring individuals over the experimental interval or could reflect that conditions within the incubators (for example, airflow patterns and/or static charges) were not conducive to airborne transfer of conidia.

The fungal skin lesions that developed in treated and contact-exposed animals were indistinguishable from those that occurred in the positive control bats (Fig. 1). Additionally, the prevalence of infection was similar between the two groups (Table 1), indicating that the treated group did not develop disease from exposure to an excessively high dose of conidia. Similar disease pathology between groups also indicates that the contact-exposed bats did not develop WNS through exposure to an agent other than *G. destructans*. Histological examination of hearts, intestines, livers, lungs and kidneys from a subset of animals (positive control group  $n = 5$ , negative control group  $n = 3$ , treated group  $n = 10$ , contact exposure group  $n = 5$ ) did not reveal any tissue damage or other signs of infectious processes that might have predisposed the animals to skin infection by *G. destructans*. Furthermore, live *G. destructans* was cultured from the skin of bats confirmed to have WNS lesions. Development of lesions diagnostic for WNS in the absence of other signs of disease provides the first experimental evidence that *G. destructans* is a primary pathogen and causes WNS in healthy bats.

The large-scale mortality seen in wild bat populations with WNS was not observed in the treated or contact exposure groups. Although all of the positive control animals died before the termination of the trial, survivorship ( $P = 0.72$ ) and body mass index (BMI;  $P = 0.96$ ) of the remaining groups did not significantly differ from the negative control group (Fig. 2a). The lack of WNS-related mortality in the treated and contact exposure groups is best explained by the short period of time these groups were exposed to *G. destructans*. On the basis of an analysis of wild bats submitted to the US Geological

<sup>1</sup>Molecular and Environmental Toxicology Center, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA. <sup>2</sup>National Wildlife Health Center, US Geological Survey, Madison, Wisconsin 53711, USA. <sup>3</sup>Wisconsin Veterinary Diagnostic Laboratory, Madison, Wisconsin 53706, USA. <sup>4</sup>Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996, USA. <sup>5</sup>Fort Collins Science Center, US Geological Survey, Fort Collins, Colorado 80526, USA. <sup>6</sup>New York Department of Environmental Conservation, Albany, New York 12233, USA. <sup>7</sup>US Fish and Wildlife Service, Hadley, Massachusetts 01035, USA. <sup>8</sup>Wisconsin Department of Natural Resources, Madison, Wisconsin 53707, USA. <sup>9</sup>Department of Biology, Bucknell University, Lewisburg, Pennsylvania 17837, USA.



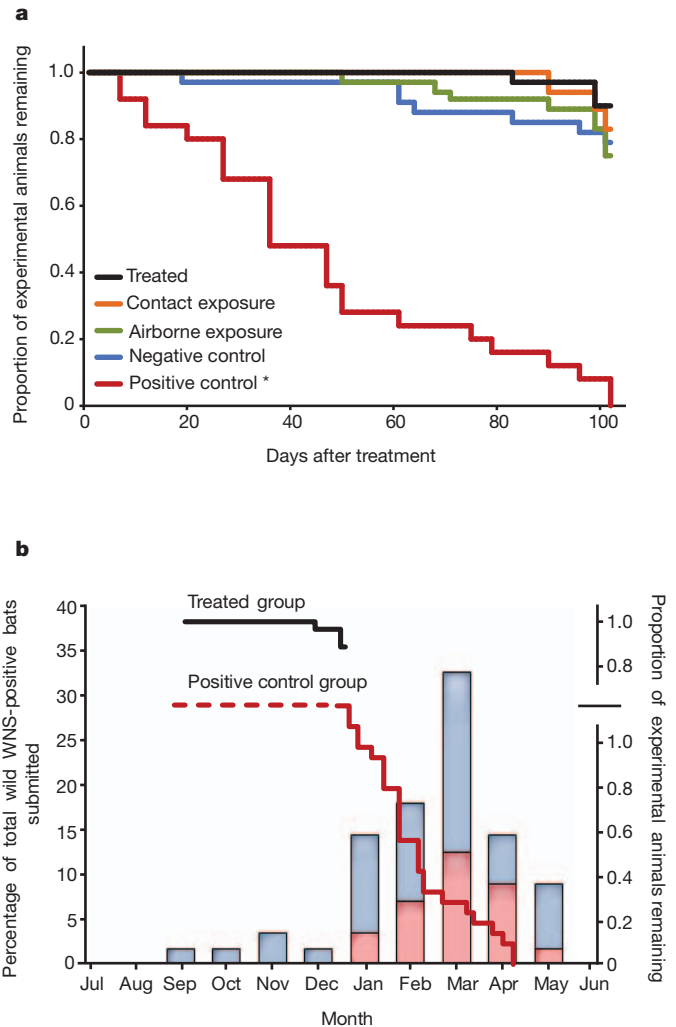
**Figure 1 | Histological sections of representative wing membranes (periodic acid-Schiff stain).** **a**, Normal wing membrane of a healthy bat from the negative control group showing no signs of fungal growth. **b**, **c**, WNS lesions, including invasion of the underlying connective tissue by fungal hyphae (arrows), are visible in sections from a bat with WNS from the positive control group (**b**) and a bat from the treated group that developed WNS after experimental exposure to *G. destructans* (**c**). Insets are higher magnification images and scale bars indicate 20 µm.

Survey (USGS)–National Wildlife Health Center (NWHC) for diagnostic testing (January 2008 to June 2011), WNS lesions have seasonally first been detected during autumn (late September), just before the start of long-term hibernation; major mortality events caused by WNS have seasonally not been observed among wild bats until the end of January (Fig. 2b). These data indicate that mortality

**Table 1 | Development of WNS in experimentally infected bats**

Treatment group	Number with WNS lesions present	Number with no WNS lesions	Total	Per cent infected
Negative control	0	34	34	0
Treated	29	0	29	100
Contact exposure	16	2	18	89
Airborne exposure	0	36	36	0
Positive control	25	0	25	100

The data show prevalence of WNS-associated fungal infections established in groups of healthy little brown bats inoculated with conidia of *G. destructans* from pure culture or exposed to bats known to have WNS (positive control group). Infection status was determined by histological examination of the wing.



**Figure 2 | Survival curves.** **a**, Survival curves for the treated ( $n = 29$ ), contact exposure ( $n = 18$ ), airborne exposure ( $n = 36$ ), negative control ( $n = 34$ ) and positive control ( $n = 25$ ) groups. Bats in the positive control group, which consisted of animals naturally infected with WNS at the time they were collected, exhibited significantly decreased survival (asterisk) relative to the other groups ( $P < 0.001$ ). Survival among bats of the remaining groups did not differ significantly from one another ( $P = 0.72$ ). **b**, Percentage of bats submitted by month (January 2008 to June 2011) to the USGS–National Wildlife Health Center that tested positive for WNS ( $n = 54$  submission events). The blue bars represent submissions that were not associated with major mortality events; the red bars depict submissions associated with high mortality. Annually, WNS-associated mortality events are first observed in January; the number of submissions involving mortality events for a given month peaks in March. Assuming the positive control bats were first exposed to *G. destructans* in late September, mortality due to WNS did not occur in the laboratory until approximately 120 days after exposure, consistent with what is observed in free-ranging wild bats (the dotted line represents the exposure period in the wild before the animals were collected for this study). The duration of this infection trial (102 days) was insufficient to observe WNS-associated mortality in the treated and contact exposure groups (the treated group mortality curve is shifted such that duration of exposure corresponds to that of the positive control group; contact and airborne exposure group mortality curves are not shown).

from WNS does not manifest until approximately 120 days after bats enter hibernation and assume a cold physiological state conducive to proliferation of *G. destructans*; mortality subsequently peaks about 180 days after bats first enter hibernacula (in the month of March). Assuming that initial exposure of positive-control bats to the fungus occurred in late September, these animals survived about 110 to 205 days after exposure, with approximately 50% having died by the

150-day mark. The treated and contact-exposure bats were only exposed to *G. destructans* for 102 days. Thus, the experiment was terminated before the disease had progressed to the degree that mortality would be expected among treated and contact-exposed animals.

Our work demonstrates experimental infection of little brown bats by *G. destructans* with subsequent development of WNS in the absence of underlying health conditions. It follows that the recent widespread detection of *G. destructans* in Europe without apparent detriment to bat populations indicates that the fungus may be endemic to that region where it co-evolved with continental bat species<sup>8,10</sup>. In North America, the data indicate that WNS originated at a single site<sup>1,18</sup> with high tourist traffic, consistent with the introduction of an exotic species<sup>19</sup>. Thus, the pathological effects caused by *G. destructans* in North American bats may reflect exposure of a naive host population to a novel pathogen. Future studies are needed to investigate the origin of *G. destructans* in North America and to elucidate differences in physiology and behaviour between North American and European bats that might account for disparate disease outcomes observed among the two continents.

Fungal pathogens have the unique capacity to drive host populations to extinction because of their ability to survive in host-free environments<sup>5</sup>. Given the high mortality rate and speed at which WNS has spread, the disease has the potential to decimate North American bat populations and cause species extinctions<sup>20</sup> similar to those documented for amphibians affected by chytridiomycosis<sup>21</sup>. Advancement of WNS research and management has been limited by uncertainty over the causative agent of this disease. With the causative agent now conclusively identified through fulfilment of Koch's postulates, future research efforts can focus on mitigating the effects of WNS before hibernating bat populations suffer losses beyond the point of recovery.

## METHODS SUMMARY

Little brown bats (*Myotis lucifugus*) naturally infected with WNS (positive control group;  $n = 25$ ) were collected from a hibernaculum in New York. Healthy (based upon body condition and histopathology findings) little brown bats were collected from a hibernaculum in Wisconsin outside of the known range of WNS. Healthy bats were divided into four groups: negative control ( $n = 34$ ), treated ( $n = 29$ ), contact exposure ( $n = 18$ ) and airborne exposure ( $n = 36$ ). Conidia of *G. destructans* ( $5 \times 10^5$  conidia suspended in 20  $\mu$ l of phosphate buffered saline solution containing 0.5% Tween 20 (PBST)) were applied to one of the wings of bats in the treated group, and an additional  $5 \times 10^5$  conidia were applied to the fur between the eye and ear. Negative control bats were treated identically with PBST lacking conidia. Animals were maintained in mesh enclosures (Supplementary Fig. 1) under conditions approximating bat hibernacula for 102 days. The experimental end point was set to correspond with the timing by which wild bats naturally emerge from hibernation. Infection status was determined by histological examination of the muzzle and skin from each wing<sup>4</sup>, and *G. destructans* was re-isolated from wing skin as previously described<sup>22</sup>. The identity of fungal isolates resembling *G. destructans* was confirmed by PCR amplification/double-stranded sequence analysis of the rRNA gene internal transcribed spacer<sup>23</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 17 August; accepted 22 September 2011.**

**Published online 26 October 2011.**

1. Blehert, D. S. *et al.* Bat white-nose syndrome: an emerging fungal pathogen? *Science* **323**, 227 (2009).
2. Turner, G. G., Reeder, D. M. & Coleman, J. T. H. A five-year assessment of mortality and geographic spread of white-nose syndrome in North American bats and a look to the future. *Bat Res. News* **52**, 13–27 (2011).

3. Gargas, A., Trest, M. T., Christensen, M., Volk, T. J. & Blehert, D. S. *Geomyces destructans* sp. nov. associated with bat white-nose syndrome. *Mycotaxon* **108**, 147–154 (2009).
4. Meteyer, C. U. *et al.* Histopathologic criteria to confirm white-nose syndrome in bats. *J. Vet. Diagn. Invest.* **21**, 411–414 (2009).
5. Casadevall, A. Fungal virulence, vertebrate endothermy, and dinosaur extinction: is there a connection? *Fungal Genet. Biol.* **42**, 98–106 (2005).
6. Blanco, J. L. & Garcia, M. E. Immune response to fungal infections. *Vet. Immunol. Immunopathol.* **125**, 47–70 (2008).
7. Maschmeyer, G., Calandra, T., Singh, N., Wiley, J. & Perfect, J. Invasive mould infections: a multi-disciplinary update. *Med. Mycol.* **47**, 571–583 (2009).
8. Wibbelt, G. *et al.* White-nose syndrome fungus (*Geomyces destructans*) in bats, Europe. *Emerg. Infect. Dis.* **16**, 1237–1242 (2010).
9. Martinková, N. *et al.* Increasing incidence of *Geomyces destructans* fungus in bats from the Czech Republic and Slovakia. *PLoS ONE* **5**, e13853 (2010).
10. Puechmaile, S. J. *et al.* Pan-European distribution of white-nose syndrome fungus (*Geomyces destructans*) not associated with mass mortality. *PLoS ONE* **6**, e19167 (2011).
11. Turner, G. G. & Reeder, D. M. Update of white-nose syndrome in bats, September 2009. *Bat Res. News* **50**, 47–53 (2009).
12. Puechmaile, S. J. *et al.* White-nose syndrome fungus (*Geomyces destructans*) in bat, France. *Emerg. Infect. Dis.* **16**, 290–293 (2010).
13. Koch, R. Die aetiologie der tuberkulose. *Mitt. Kaiser. Gesundh.* **2**, 1–88 (1884).
14. Cryan, P. M., Meteyer, C. U., Boyles, J. G. & Blehert, D. S. Wing pathology of white-nose syndrome in bats suggests life-threatening disruption of physiology. *BMC Biol.* **8**, 135 (2010).
15. Foley, J., Clifford, D., Castle, K., Cryan, P. M. & Ostfeld, R. S. Investigating and managing the rapid emergence of white-nose syndrome, a novel, fatal, infectious disease of hibernating bats. *Conserv. Biol.* **25**, 223–231 (2011).
16. Thomas, D. W., Fenton, M. B. & Barclay, M. R. Social behavior of the little brown bat, *Myotis lucifugus*. *Behav. Ecol. Sociobiol.* **6**, 129–136 (1979).
17. Fenton, M. B. Summer activity of *Myotis lucifugus* (Chiroptera: Vespertilionidae) at hibernacula in Ontario and Quebec. *Can. J. Zool.* **47**, 597–602 (1969).
18. Rajkumar, S. S. *et al.* Clonal genotype of *Geomyces destructans* among bats with white-nose syndrome, New York, USA. *Emerg. Infect. Dis.* **17**, 1273–1276 (2011).
19. Desprez-Loustau, M. L. *et al.* The fungal dimension of biological invasions. *Trends Ecol. Evol.* **22**, 472–480 (2007).
20. Frick, W. F. *et al.* An emerging disease causes regional population collapse of a common North American bat species. *Science* **329**, 679–682 (2010).
21. Skerratt, L. F. *et al.* Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth* **4**, 125–134 (2007).
22. Lorch, J. M. *et al.* Rapid polymerase chain reaction diagnosis of white-nose syndrome in bats. *J. Vet. Diagn. Invest.* **22**, 224–230 (2010).
23. White, T. J., Bruns, T., Lee, S. & Taylor, J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR Protocols: A Guide to Methods and Applications* 315–322 (Academic Press, 1990).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** Financial support for this project was provided by the US Geological Survey, the US Fish and Wildlife Service, Bat Conservation International, and the Indiana State University Center for North American Bat Research and Conservation. Use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government. We acknowledge E. Buckles, S. Darling and T. Tomasi for discussions during the development of this project. We thank N. Keller for comments during the preparation of this manuscript. We also thank J. P. White, R. Dusek, A. Klein, L. Leppert, K. Schuler, C. L. White and NWHC Animal Care Staff for their help with set-up and maintenance of animals, and D. Berndt, L. Muller, A. Pike and M. Verant for their assistance during necropsies. We thank X. Yu for help with statistical analyses.

**Author Contributions** D.S.B., A.C.H., P.M.C. and D.M.R. designed the study. A.C.H., J.T.H.C. and D.N.R. collected wild animals for the study. J.M.L., D.S.B., C.U.M., M.J.B., J.G.B., A.C.H., A.E.B., J.T.H.C., D.N.R. and D.M.R. performed the experiment and/or assisted to collect samples upon completion. C.U.M. and M.J.B. read and interpreted histopathology. J.M.L. and J.G.B. analysed the experimental data. A.E.B. compiled data on wild bats submitted to the NWHC for diagnostic testing. J.M.L. wrote the manuscript and all co-authors provided input. D.S.B. supervised data analyses and edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.S.B. ([dblehert@usgs.gov](mailto:dblehert@usgs.gov)).

## METHODS

**Animals.** This study was conducted at the NWHC in accordance with Institutional Animal Care and Use Committee Experimental Protocol 081118. WNS-positive little brown bats (*Myotis lucifugus*) (positive control group;  $n = 25$ ) were collected from a hibernaculum in New York in January 2009; only bats showing visible signs of fungal growth on the muzzle and/or wings were collected from the New York site. Healthy (based upon body condition and histopathology findings) little brown bats were collected from a hibernaculum in Wisconsin (approximately 1,000 km distant from the known range of WNS at the time that animals were collected). Bats were transported to the NWHC in coolers at approximately 7 °C.

**Experimental infection.** Healthy bats were randomly (except for ensuring nearly equal sex ratios) divided into groups: negative control ( $n = 34$ ), treated ( $n = 29$ ), contact exposure ( $n = 18$ ) and airborne exposure ( $n = 36$ ). Negative control, positive control and treated groups were maintained in separate rooms. Animals in the contact exposure group were placed in the same enclosures as the positive control group. Animals in the airborne exposure group were split evenly between separate enclosures, each located 1.3 cm from enclosures housing the positive control and treated groups (Supplementary Fig. 1).

Conidia were harvested from 60-day-old cultures of the type strain of *G. destructans*<sup>3</sup> (American Type Culture Collection number ATCC MYA-4855) by flooding plates with phosphate buffered saline solution containing 0.5% Tween20 (PBST). Conidia were washed, enumerated and re-suspended in PBST. Twenty microlitres of the conidial suspension containing  $5 \times 10^5$  conidia were pipetted directly onto the dorsal surface of one of the wings of bats in the treated group; an additional aliquot (20 µl) was pipetted onto the fur between the eye and ear. Negative control bats were treated identically with PBST lacking conidia. Bats were housed in mesh enclosures (Reptaria; Apogee) within refrigerators (SRC Refrigeration) under conditions approximating bat hibernacula (complete darkness, approximately 6.5 °C and 82% relative humidity) for 102 days. Termination of the experiment corresponded to the time period during which wild bats begin to emerge from hibernation. Temperatures were recorded daily in each refrigerator to ensure that appropriate hibernation conditions were maintained. The mean

temperatures ( $\pm$  standard deviation) for the refrigerators were as follows: negative control group,  $6.4 \pm 0.8$  °C; positive control, airborne exposure (in part) and contact exposure groups,  $6.7 \pm 0.4$  °C; treated and airborne (in part) exposure groups,  $6.4 \pm 0.8$  °C. BMI was calculated by dividing body mass at the time that the bats were euthanized by forearm length. Because animals that died naturally during the trial became desiccated, BMI was only calculated for bats that were euthanized.

Diagnosis of WNS was made through histological examination of the muzzle and a portion of skin from each wing<sup>4</sup>. *G. destructans* was re-isolated in culture from wing skin as described previously<sup>22</sup> and identified by PCR amplification/double-stranded sequence analysis of the rRNA gene internal transcribed spacer<sup>23</sup>.

**Statistical analyses.** Survivorship was compared among groups using the Gehan–Breslow survival test (SigmaPlot 11.0; Systat Software) because this method gives more weight to animals that died naturally during the experiment and less weight to the large number of censored data points (that is, euthanized animals) at the end of the experiment. Pair-wise comparisons were examined with the Holm–Sidak procedure (significance at  $P < 0.05$ ). BMI was compared among groups (negative control group,  $n = 27$ ; treated group,  $n = 25$ ; contact exposure group,  $n = 15$ ; airborne exposure group,  $n = 27$ ) using an analysis of variance test (significance at  $P < 0.05$ ) after confirming that the data met assumptions of normality (Shapiro–Wilk test,  $P = 0.07$ ) and equal variances (Levene median test,  $P = 0.87$ ). One bat from the treated group was excluded from the BMI analysis because its weight was not measured before euthanasia and sample collection. Three bats from the treated group were euthanized 34 days after exposure to assess whether WNS lesions were developing; WNS lesions were not detected in these animals. Because these three animals were prematurely removed from the experiment, they were excluded from further analyses and are not represented in the specified sample sizes.

**Equipment and settings.** Prepared tissue sections were examined using an Olympus BH-2 upright microscope with SPlan Apo  $\times 40$  and  $\times 100$  objectives (Olympus Optical). Images were collected in tagged image file format using a digital colour camera (Insight2) and Spot Basic Version 4.0.8 (Diagnostic Instruments).